

31

Journal of PROJECTIVE TECHNIQUES

PUBLISHED BY THE SOCIETY FOR PROJECTIVE TECHNIQUES, INC.

TABLE OF CONTENTS

A Guide for Research Use of the Blacky Pictures.....	GERALD S. BLUM	3
The Effect of Color on Productivity in Cards VIII, IX, X of the Rorschach.....	JOHN E. EXNER, JR.	30
Levels of Awareness in Projective Tests GORDON FILMER-BENNETT and WALTER G. KLOPPER		34
Factors Influencing Rotation in the Bender-Gestalt Performance of Children.....	JERRY B. FULLER and GILLES CHAGNON	36
Responses of Children to a Group Projective Film and to the Ror- schach, CAT, Despert Fables and D-A-P	MARY E. HAWORTH	47
Stylistic Consistency Among Aesthetic Preferences ROBERT H. KNAPP and HELEN EHLINGER		61
The Utilization of the Psychologist-Patient Relationship in Diagnostic Testing.....	THEODORE LEVENTHAL, HOWARD J. SLEPIAN, MARTIN R. GLUCK and BERNARD P. ROSENBLATT	66
Interlevel Disparity and Predictive Efficiency.....	JAMES C. McGREEVEY	80
An Indirect Validation Study of the Draw-A-Person Test Through the Cartoons of William Steig.....	DONALD McGUIRL and C. SCOTT MOSS	88
The Role of the TAT in the Measurement of Achievement as a Function of Expectancy.....	BERNARD I. MURSTEIN and HERBERT L. COLLIER	96
Projective Techniques in the Selection of Management Personnel.....	JOSEPH G. PHELAN	102
Intraindividual Response Variability in Figure Drawing Tasks.....	JOEL SHANAN	105
The Natural History of an Outcome Prediction.....	RICHARD S. SIEGAL, IRWIN C. ROSEN and GERALD A. EHRENREICH	112
Manifestations of Primary Process Thinking in Schizophrenia.....	LLOYD H. SILVERMAN, BENJAMIN LAPKIN and IRA S. ROSENBAUM	117
Book Reviews		128
Announcements		134
Directory		135



OFFICERS OF THE SOCIETY FOR PROJECTIVE TECHNIQUES, Inc.

<i>President</i>	EDWIN S. SHNEIDMAN
<i>President-Elect</i>	PAULINE G. VORHAUS
<i>Past President</i>	BERTRAM R. FORER
<i>Eastern Representative</i>	L. JOSEPH STONE
<i>Western Representative</i>	EVELYN HOOKER
<i>Secretary</i>	WALTER G. KLOPPER
<i>Treasurer</i>	HARRY V. MCNEILL
<i>Editor</i>	BRUNO KLOPPER
<i>Executive Editor</i>	BERTRAM R. FORER
<i>Executive Secretary</i>	GORDON F. DERNER

ADVISORY COMMITTEE

FRANKLIN G. EBAUGH
 LAWRENCE K. FRANK
 BRUNO KLOPPER, *Chairman*
 NOLAN D. C. LEWIS
 MORTIMER M. MEYER
 GARDNER MURPHY
 DAVID VORHAUS

JOURNAL OF PROJECTIVE TECHNIQUES, published quarterly by the Society for Projective Techniques, Inc. Annual subscription (calendar year) \$8.00, foreign \$8.50. Single copies \$2.00. Publication office: 210 E. Wilson Ave., Glendale 6, Calif. Subscriptions, changes of address and business matters should be addressed to the Society for Projective Techniques, Inc., 210 E. Wilson Ave., Glendale 6, Calif., or Adelphi College, Garden City, N.Y. Address all Editorial Matters to: Bertram R. Forer, Suite 307, 8833 Sunset Blvd., Los Angeles 69, Calif. Entered as 2nd class matter at Glendale, Calif.

Copyright 1962 by the Society for Projective Techniques, Inc.

JOURNAL OF PROJECTIVE TECHNIQUES

VOL. 26

MARCH, 1962

No. 1

. . .

EDITOR

BRUNO KLOPPER

EXECUTIVE EDITOR

BERTRAM R. FORER

ASSOCIATE EXECUTIVE EDITORS

WALTHER JOEL
JEANNE M. REITZEL
EDWIN S. SHNEIDMAN

STAFF EDITORS

EVELYN CRUMPTON
FLORENCE DIAMOND
NORMAN L. FARBEROW
HERMAN FEIFEL
FRED J. GOLDSTEIN
JAMES HOLT

WALTER G. KLOPPER
LEONARD KOROT
LEONARD B. OLINGER
NICHOLAS ROSE
JOHN WEIR
IRLA LEE ZIMMERMAN

A Guide for Research Use of the Blacky Pictures

GERALD S. BLUM
University of Michigan

The attainment of a sound scoring system for a projective technique does not come easily. Indulging one's own fantasy along these lines, it would be nice to administer the test to large samples of normal and pathological individuals varying in age, sex, etc.; then derive empirically the most meaningful set of categories encompassing all their responses; and finally relate scores on those categories to a host of criterion variables. The pursuit of such a program in Blacky's behalf seemed quite appropriate. Though not a gay dog, he (she) is obviously addicted to dreams of glory. What could be finer than to be outfitted in a shiny new scoring system?

In reality, though, Blacky is a creature born to compromise and the plans for his refurbishing likewise had to be cut down to size. We restricted the sample to a couple of hundred male undergraduates previously tested in groups at an eastern, a midwestern, or a far western university. Available kinds of criteria numbered a few dozen but only for a segment of the total sample. Within these limits, though, an exhaustive study of all responses was conducted.¹ (Parenthetically, it can be added that the analysis took several years and a large sum of money,² so the writer is

not overly defensive about the above restrictions!)

Since their inception in 1946 the Blacky Pictures (Blum, 1950) have been used to assess intensity of disturbance or conflict along a series of dimensions derived from psychoanalytic theory: oral eroticism, oral sadism, anal expulsiveness and retentiveness, oedipal intensity, masturbation guilt, castration anxiety (males) or penis envy (females), identification process, sibling rivalry, guilt feelings, ego ideal, and narcissistic or anacletic love object. The original scoring system (Blum, 1949) was constructed in a *priori* fashion from four sources of data relevant to each dimension. These included spontaneous stories given to the cartoons, structured inquiry items, cartoon preferences, and related comments cropping up on cartoons other than the one under consideration. In 1951 an internal consistency analysis of several hundred protocols led to a revised set of procedures (Blum, 1951) which has been employed from that time on. By 1958, however, the accessibility of an electronic computer had instigated a nagging thought that the grandiose approach might not be so fantastic after all. The present report describes the venture which ensued.

PROCEDURES

Derivation of new scoring system

The empirical analysis was carried out separately on a picture-by-picture basis. For each of the eleven cartoons all spontaneous story themes were coded from protocols of the 210 Ss; inquiry item answers entered; expressions of "like" or "dislike" (plus "best" or "worst" judgment if assigned to that particular cartoon) recorded; and related comments gleaned from responses to other cartoons.

¹ Grateful acknowledgment is hereby extended to the following individuals: development of the scoring categories was carried out with the assistance of Myrtle Blum and Pearl Franzblau; Abram Minkowich and Mary Lee Pierce prepared the materials for statistical analysis; and Lillian Kelly processed the data through the various phases of computer operation. E. Lowell Kelly and William Hays furnished helpful consultation on methodology; Linda Shure and William Winter offered suggestions for improving the manuscript.

² Financial support came partly from National Institute of Mental Health grant M-1286 and a Ford Foundation Mental Health grant.

The correlational program in use for the IBM 650 computer at the time dictated an upper limit of 38 variables in the matrix for a picture. Highly infrequent responses (given by fewer than 5 of the 210 Ss) were therefore dropped. On most pictures it was also necessary to combine some closely allied themes into a single category in order to remain within the limit.

Once the 38 variables for a cartoon were selected and their presence or absence ascertained for all Ss, their intercorrelations were computed. Next an orthogonal factor analysis was performed upon the matrix using the centroid method of extraction and varimax rotation procedures. A rather arbitrary decision, guided by the number of dimensions per picture in the earlier scoring systems, was made to extract four factors from each matrix, allowing an opportunity for somewhat more than the expected number to appear. Most often three factors per picture were able to be identified with some confidence as meaningful statistical and psychological entities. Overall 30 such factors emerged, three on eight of the cartoons, and two each on the three others.

Positive factor loadings determined the choice of variables to include within a specific factor score. In rare instances a variable loaded almost equally on two different factors and was therefore retained on both. The modal number of variables per factor is 7, with a range of 5-10. Every S was then given a numerical score on each of the 30 factors, depending upon how many of the relevant variables for a factor were present in his protocol. Although the magnitude of factor loadings varied considerably, the lack of independence among some of the variables was felt to be a sufficiently distorting element to preclude differential weights. Accordingly one point is given for each factor variable and a factor score is simply the sum of those variables

present in an individual's record.

Finally the numerical factor scores were themselves intercorrelated *across* pictures and attempts subsequently made to discover additional factors beyond the intra-picture ones already extracted. These latter efforts have supported a conclusion that the 30 factors do *not* lend themselves to broad clustering but instead maintain their specificity. (Correlations between pairs of variables will be given later along with other findings.)

Available criterion variables

As discussed in an early article on the validity of the Blacky Pictures (Blum & Hunt, 1952) truly definitive criterion variables do not exist. Instead it is necessary to pursue a construct validity approach by relating Blacky scores to a wide assortment of dependent variables in the hope of establishing a meaningful frame of reference. The present study makes use of a great deal of supplementary information previously collected from certain members of the sample. For 155 Ss sibling data—including number, ordinal position, and sex—were available. Major field of specialization in college was known for 71 Ss. Forty-four members of a social fraternity had participated in an extensive research program³ of a year's duration and it was from this group that the bulk of dependent variables were derived. The latter are listed below:⁴

Strong Vocational Interest Blank (SVI)

- Group III: Production Manager
- Group V: Social Service
- Group VI: Musician

³ This project was supported by NIMH grant M-659. In addition to the writer as principal investigator, participants included Richard Blake, Arthur Cohen, Sherman Nelson, Leon Sirota, Justin Weiss, and William Winter.

⁴ Detailed descriptions of the data collection procedures are given elsewhere (Sirota, 1957). Certain variables were subsequently excluded because of lack of spread among responses or incompleteness of information for the group.

Group VIII: Business Detail
 Group IX: Business Contact
 Group X: Verbal
 Group XI: President of Manufacturing Co.
 Interest Maturity
 Occupational Level
 Masculinity of Interests

*Allport-Vernon-Lindzey
 Study of Values (AVL)*

Theoretical
 Economic
 Aesthetic
 Political

*Blacky Defense Preference Inquiry
 (DPI) (Blum, 1956)*

Avoidance	} measured on each of the 11 pictures
Reaction Formation	
Projection	
Regression	
Intellectualization	

Biographical Questionnaire

Parents native or foreign born
 Size of town in which raised
 Annual family income
 Reported strictness of childhood discipline
 Reported amount of punishment in childhood
 Father or mother perceived as final authority in discipline
 Characteristic modes of discipline employed by father and mother
 Reported amount of friction between parents
 Religious training
 Personal habits, including smoking and drinking
 Checklist of past and present physical complaints
 Extracurricular activities
 Grade-point average in college

Miscellaneous Measures

Perceptual accuracy: ability to discriminate Blacky pictures presented tachistoscopically (derived from an experimental study of perceptual vigilance and defense).

Social perception: based on coding of descriptive data provided by

Ss about their perception of the fraternity group in which they lived.

Oral eroticism (supportive interpersonal relationships stressed; oral gratification mentioned)

Sibling hostility (emphasis on laziness or lack of responsibility of members)

Ego ideal (group's incentives, achievement orientation)

Intellectual orientation (group viewed in terms of academic standards)

Interpersonal mechanisms manifested: based on responses to a story completion test involving situations related to sibling rivalry, oedipal intensity, oral eroticism, and oral sadism; scoring categories adapted from Freedman, et al (1951)

Dominating
 Proud
 Rejecting
 Punitive
 Respectful
 Loving
 Supportive

Humor preference and recall: based on a specially devised humor test and subsequent recall questionnaire.

Reactions to cartoons and jokes with psychosexual content

Recall of cartoons and jokes with psychosexual content

FINDINGS

In this section each of the 30 obtained Blacky factors will be taken up in turn. First the test variables comprising the factor score will be listed; next those statistically significant positive and negative relationships of the factor to other factors, defense preferences, and the remaining criterion variables will be given;⁵

⁵ The reported relationships are all significant at the .05 level (two-tailed test) or better. For the factor intercorrelations

and finally an attempt will be made to interpret the psychological meaning of the factor in light of the evidence. (Roman numerals in the factor and DPI designations refer to the Blacky pictures from which they are derived; capital letters differentiate the several orthogonal factors based on the same picture.)

ORAL EROTICISM (*Cartoon I*)

Factor I-A. Oral Craving

Inquiry:

- 3b B would rather stay as long as possible to be sure he gets enough nourishment.
- 4a B is a little glutton who never stops eating.
- 5b B will want to continue being nursed until he's quite a bit older.
- 6a B will rather eat than do most anything else when he grows older.

Theme:

B is too old to be nursing; too large.

Related comments:

Food, water, hunger, containers, etc. (Note: related comments are not scored on Cartoon I; bone is not scored in this category; one point is given for each cartoon having related comments but no more than one point can be assigned to any one

($N=210$) the magnitude had to reach .139 or higher; for the sibling data ($N=155$) .161; for field of specialization ($N=71$) the corresponding r was .235; and for all other variables at least .297. It is impossible of course to rule out those relationships achieving statistical significance on a chance basis alone, so that a given list is viewed broadly as a constellation providing the currently best frame of reference for interpreting the factor. Before lending special emphasis to any single correlation, it should be subjected to cross-validation. A reader wishing detailed statistical information, omitted from this presentation because of space requirements, is invited to write directly to the author.

cartoon even if it has two or more comments.⁶)

Positive relationships

- Supply-seeking (Factor II-B)
- Resentment over oral deprivation (Factor II-C)
- Overt hostility toward sibling and mother (Factor VIII-A)
- Negative perception of self and father (Factor X-B)
- Intellectualization (DPI-V)
- Father reported to use deprivation as punishment

Negative relationships

- Playfulness (Factor II-A)
- Regression (DPI-VII)
- Avoidance (DPI-X)
- Father reported as using corporal punishment

Interpretation.

From the above relationships a hypothetical person scoring high on this factor would be expected to possess an intense, almost voracious craving for oral supplies, in the psychoanalytic sense of the term. The craving can never be fully satisfied and there is a perpetual concern over potential deprivation. Reactions to lack of gratification are characteristically resentful, with hostility directed toward mother primarily and toward competing sibs. Father too is seen as depriving and does not serve as an adequate ego ideal figure. Wish-fulfilling fantasies, although prevalent, do not obscure a basically pessimistic attitude in regard to the attainment of happiness. No consistent pattern of defense preferences is revealed.

Factor I-B. Oral Rejection

Inquiry:

- 1b B is unhappy.
- 2c M feels rather unhappy.
- 4c B sometimes doesn't get enough to replace all the energy he burns up.

⁶ As mentioned earlier, the factor score is simply the sum of the number of points accrued from inquiry items, cartoon preferences, related comments, and themes.

- 5c B feels M would like to turn him loose right now.
 6c B will never get enough to satisfy his appetite.

Theme:

M rejecting or strongly disinterested.

(Examples: "disgusted"—"tired"—"petered out"—"reluctant to feed B"—"wants to get away"—"all played out"⁷ [more than "not paying attention" or "only doing duty"])

Positive relationships

Resentment over oral deprivation (Factor II-C)

Negative perception of self and father (Factor X-B)

Narcissism (Factor XI-C)

Youngest of three or more siblings

Major in business or pre-law

Poor eyesight

Perceptual accuracy

Large amount of punishment in childhood reported

Dislike for oedipal jokes (Humor)

Supportive (Interpersonal Mechanisms)

Negative relationships

Playfulness (Factor II-A)

Oldest of three or more siblings

Liberal arts major

Excessive smoking

Reaction formation (DPI-VII)

Interpretation.

The constellation here is clearly one of maternal rejection. Dependency needs are thwarted by the mother, who apparently is unable to bestow affection upon the youngest of several children. Frequent punishment is reported as having occurred during childhood. This depressing, strained relationship between mother and son appears to have severe repercussions in the latter's heterosexual

adjustment, reflected by a highly narcissistic approach to the choice of a love object. Dislike for jokes of an oedipal nature also points to the origin of heterosexual difficulties in the unsatisfactory maternal relationship. The inadequacy of father as an ego ideal figure further strengthens the tendency to turn away from the family and toward others in the search for narcissistic interpersonal gratification. Vocational preferences expressed by field of specialization involve typically well-defined interactions and the evidence suggests that a supportive, friendly role can probably be played in such contexts. The perceptual data, in which poor eyesight paradoxically is accompanied by greater accuracy in discriminating Blacky pictures flashed tachistoscopically (glasses worn during experiment of course), possibly indicate a special effort to be alert and responsive to external stimulation.

Factor I-C. Sugar-Coating

Inquiry:

1a B is happy.

2a M feels very contented.

6a B will rather eat than do most anything else.

Preferences:

Like } (one point for each)
 Best }

Theme:

M contented or enjoying herself while feeding B.

(Example: "M quite happy"—"M very cooperative")

Positive relationships

Overwhelming castration conflict (Factor VI-A)

Overtly positive perception of self and father (Factor X-A)

Negative relationships

Fear of punishment for masturbation (Factor V-A)

Major in business or pre-law

Business contact (SVI)

Dislike for oedipal jokes (Humor)

⁷ A note of caution must be added concerning the scoring of open-ended or spontaneous material. Although the cited examples are merely illustrative, they should serve as firm guides to the interpretation of the category itself. In other words, the scorer is not permitted to depart from the literal intent of the category.

Supportive (Interpersonal Mechanisms)

Interpretation.

This factor is the first of several which reflect a defensive style of responding to the cartoons. By proclaiming that everything is extremely fine, the individual does not have to come to grips with the problem area itself. Supporting such an interpretation is the positive correlation with Factor X-A, which presents both Blacky and Papa in a highly favorable light, and the negative one with Factor V-A, where fears and misgivings are openly expressed. A suggested motive for the defense is the existence of overwhelming castration conflict.

ORAL SADISM (Cartoon II)

Factor II-A. Playfulness

Inquiry:

- 3c B acts like this when he can't get enough recreation.
4a B will get tired of the collar and leave it on the ground.

Preferences:

Like } (one point for each)
Best }

Themes (including item 1 of inquiry)

- (1) B playing with collar.

(Examples: "acting tough"—"playing"—"pretending to fight collar"—"pretending collar is enemy or beast")

- (2) B exercising body, sharpening teeth, expending energy.

Positive relationships

Choosing obvious neutral responses (Factor III-B)

Reaction formation to sibling rivalry (Factor VIII-B)

Avoidance (DPI-I)

Avoidance (DPI-II)

Avoidance (DPI-IV)

Avoidance (DPI-IX)

Avoidance (DPI-X)

Major in engineering

Rejecting (Interpersonal Mechanisms)

Negative relationships

Oral craving (Factor I-A)

Oral rejection (Factor I-B)

Overt hostility toward sibling and mother (Factor VIII-A)

Negative perception of self and father (Factor X-B)

Regression (DPI-I)

Regression (DPI-II)

Intellectualization (DPI-VIII)

Dislike for castration cartoons (Humor)

Major in business or pre-law

Interpretation.

Another defensive factor, *Playfulness* serves to avoid expression of hostility toward the mother by structuring the picture in an innocuous fashion. This evasive aspect in the handling of interpersonal aggression is strikingly borne out by the repeated choice of avoidance over expressive alternatives in the DPI;⁸ positive correlations with defensive factors from other pictures; negative correlations with expressive factors; and even the selection of a field of specialization, engineering, which minimizes the necessity to deal with people. When interpersonal relations cannot be averted, reactions are likely to be rejecting and hostile.

Factor II-B. Supply-Seeking

Inquiry:

- 2c B feels like acting up this way very often.
3a B most often acts like this when he can't get enough attention.
4b B will return the collar to M.
5a If M comes on the scene, she will feed B again.
6b If M did come over to feed B, he'd put down the collar and start eating.

Theme (including item 1 of inquiry):

⁸ Avoidance and reaction formation are classed as avoidant defenses; projection, regression, and intellectualization (all of which permit some degree of expression of the unacceptable impulse) are classed as expressive ones.

B feels deserted by M.

(Examples: "B is angered by M's desertion"—"M has gone off and B wants her"—"M died and left him"—"B has a left out and lonely feeling")

Related comment:

Stealing food or bone. (Not scored on Cartoon I; one point per cartoon.)

Positive relationships

Oral craving (Factor I-A)

Concern over sexual maturation (Factor V-B)

Overt hostility toward sibling and mother (Factor VIII-A)

Negative perception of self and father (Factor X-B)

Projection (DPI-IV)

Regression (DPI-VI)

Reaction formation (DPI-VII)

Respectful (Interpersonal Mechanisms)

Mother reported as scolding

Negative relationships

Back pain

Food allergy

Childhood discipline reported as strict

Recall of oral sadistic cartoons (Humor)

Interpretation.

This factor has some overlap with *Oral Craving* (Factor I-A) but the quest for supplies seems to transcend a narrow preoccupation with food, though the latter is manifested. Desire for attention is salient and the orientation is more toward gaining approval from others. Although frustration is sometimes accompanied by hostile expression, a passive approach is quickly adopted if there is any hope of obtaining the sought after supplies. Interpersonal behavior is characteristically compliant and respectful, with the underlying passivity leading to feelings of inadequacy and sexual confusion. Both parents tend to be viewed ambivalently, at times seeming helpful and sympathetic but not enough so to obviate the necessity

to continue seeking reassurance. The defense preferences also indicate a need for greater family acceptance and close companionship. Even hostile humor has a threatening quality and tends to be easily repressed.

Factor II-C. Resentment over oral deprivation

Inquiry:

3b B most often acts like this when he can't get enough milk.

Preference:

Worst

Themes (including item 1 of inquiry):

- | | | |
|-------------------------|---|--|
| (one point
for each) | } | (1) Feeding reference implying deprivation. |
| | | (2) B is mean; uncontrollable rage. |
| | | (3) Violence or physical struggle involving M. |

(Examples of (3) "B has just killed M" — "B fought with M" — "He wrenched the collar from her neck" — "Someone else took M away by force")

Positive relationships

Oral craving (Factor I-A)

Oral rejection (Factor I-B)

Undisguised oedipal involvement (Factor IV-A)

Overt hostility toward sibling and mother (Factor VIII-A)

Mother-surrogate as love object (Factor XI-A)

Regression (DPI-II)

Regression (DPI-X)

Reaction formation (DPI-IV)

Intellectualization (DPI-X)

Major in physical or natural science

Negative relationships

Avoidance (DPI-I)

Avoidance (DPI-X)

Recall of oral sadistic cartoons (Humor)

Major in engineering

Interpretation.

An individual scoring high on this factor harbors deep-seated, almost

vicious resentment toward his mother for her neglect of his oral needs. The Blacky scene is very upsetting and provokes overtly hostile responses, manifested again on the sibling rivalry picture. Antecedent connections to Factors I-A and I-B point to the primarily oral nature of the struggle. Attempts to handle the oral sadistic conflict are regressive, with the "temper tantrum" DPI alternative being preferred. As in the previous factor, oral sadistic humor tends not to be recalled. Very striking is the concomitant oedipal involvement with the mother, for whom there is a largely undisguised sexual longing. The search for a female love object in the mother's image confirms the inability to free himself from the intensely hostile yet tempting relationship.

ANAL SADISM (Cartoon III)

Factor III-A. Exploitation

Inquiry:

- 1d B defecated there because he wanted to keep his own area neat and clean.
- 2c B is most concerned here with getting rid of his anger.
- 4d B thinks he's got M and P right where he wants them.
- 5 M will scold B (not bark at him or punish physically).
- 6 P will scold B (not bark at him or punish physically)

Preference:

Like

Theme:

B doesn't want to dirty own place, smart to go elsewhere.

(Examples: "B has sense enough not to relieve himself outside his own house. He has learned better"—"B's pretty wise not messing up ground around his own house. Smart dog!"—"B being a clever dog has chosen a spot far removed from his own kennel to relieve himself."—"B is a very clean dog but has selfish habit of

leaving his stools near his parents' doghouses")

Positive relationships

- Qualification of pervasive guilt (Factor IX-C)
- Negative perception of self and father (Factor X-B)
- Sole sibling is an older sister
- Poor eyesight
- Father reported as using corporal punishment

Negative relationships

- Denial of masturbation guilt (Factor V-C)
- Overtly positive perception of self and father (Factor X-A)
- Reaction formation (DPI-III)
- Reaction formation (DPI-VII)
- Father reported as lecturing
- Sole sibling is an older brother
- Major in engineering
- Aesthetic (AVL)
- Recall of oral sadistic jokes (Humor)

Interpretation.

The prominent feature here is the tendency toward selfish expression of aggressive impulses, especially anal exploitation. Both expulsive (venting anger) and retentive (keeping clean) anal components are manifested. The absence of aesthetic values is also consonant with such anal conflicts. A generally expressive style shows up in the deprecation of self and father, the admission of severe guilt, and negative relationships with factors and defenses primarily avoidant in nature. Of particular interest is the finding that anal exploitation occurs more in families where the only sibling is an older sister and less where there is an older brother.

Factor III-B. Choosing Obvious Neutral Responses

Inquiry:

- 1c B picked the spot by accident.
- 2b B is most concerned with relieving himself so that his system feels more comfortable.
- 3c B is automatically doing what he's been taught.

- 4a B feels that by relieving himself in the way he's been taught, he now has an opportunity to show his family what a good dog he can be
- 4c B is very happy to have control of himself
- 5 M will say nothing
- 6 P will say nothing, or will let M handle the situation

Positive relationships

- Playfulness (Factor II-A)
- Denial of masturbation guilt (Factor V-C)
- Evasion of identification issue (Factor VII-C)
- Reaction formation to sibling rivalry (Factor VIII-B)
- Intellectualization (DPI-IV)
- Avoidance (DPI-VIII)
- Father reported as lecturing

Negative relationships

- Fear of punishment for masturbation (Factor V-A)
- Father as identification object (Factor VII-A)
- Overt hostility toward sibling and mother (Factor VIII-A)
- Qualification of pervasive guilt (Factor IX-C)
- Negative perception of self and father (Factor X-B)
- Regression (DPI-IV)
- Intellectualization (DPI-VIII)
- Sole sibling is an older sister
- Intellectual orientation (Social Perception)
- Father reported as scolding
- Thumbsucking
- Dislike for castration jokes (Humor)
- Recall of oedipal cartoons (Humor)

Interpretation.

This factor represents a purely defensive effort not to give oneself away in responding to the test. The scoring variables all come from the inquiry, which is the source most susceptible to simple evasion. Virtually all the concomitant relationships support a strictly defensive interpretation, that

is, evasive responses are continually wrought out to the exclusion of any form of expressive involvement.

Factor III-C Attempted Denial of Anal Preoccupation

Inquiry

- 1a B cannot or keeps his own area neat and clean
- 1a B wants to make as little mess as possible
- 5 M will bark (not scold).

Theme

Complete omission or denial of anal reference

(Examples: "B buried a bone between one of the huts the day before and now he can't seem to remember exactly where"—"Here B is digging to bury the collar"—"Here he is digging, using up energy and attracting M"—"B seems to be kicking up some dirt. This is to release some energy he's stored up")

Related comments:

Later references to Cartoon III, defecation, or the anal region.

(Examples: "B looks on with certain misgivings for he has a faint suspicion that he will also get it in the end" (Cartoon VI)—"B has just taken his morning's morning in front of his parents' houses again. He knows he shouldn't have and P is out to get him and beat hell out of him")

(IX) "Maybe they have something to do with how I defecate, don't know what these fool things are" (VI) "B's conscience is bothering him for his defecating on the floor of his house" (IX) (One point per cartoon)

Positive relationships

- Minimizing castration anxiety (Factor VI-B)
- Avoidance (DPI-IX)
- Intellectualization (DPI-X)
- Speech difficulty
- Chronic sinus trouble
- Interest maturity (SVI)

Negative relationships

- Rejecting (Interpersonal Mechanisms)

Interpretation.

Although there is again a defensive quality as in the previous factor, the pattern is much more complex. Evidence suggests a pervasive concern over analty. At best the defense is a retentive reaction formation emphasizing compulsive cleanliness; at worst it degenerates into a primitive denial of real world percepts. Repression of guilt feelings and castration fears are also indicated. This always-threatening breakthrough of unacceptable anal impulses conforms to the psychoanalytic account of speech difficulties, which turn out to be correlated with the factor.

OEDIPAL INTENSITY (*Cartoon IV*)

Factor IV-A. Undisguised Oedipal Involvement

Inquiry:

- 3a B is most unhappy over P keeping M all to himself.
- 4c B suspects P is having his own way about things.
- 5 P will punish B physically
(Examples: "Beat him—" "Spank him" — "Bite him" — "Kick him")

- 7 (Answer to question "why?" in conjunction with choice of alternative b): Sexual aspects of B's relationship to M.

(Examples: "Then P would be able to observe the up-to-date approved methods of lovemaking" — "incest" — "B would be having the fun and I'm all for B" — "M and son neck")

Themes (including item 1 of inquiry):

- (1) B competitive toward P.

(Examples: "B wants to make love feels hostility toward P"—"B wants to be in P's place"—"P is B's rival"—"P is inferior to B")

- (2) B wants to love M.

(Examples: "B wants to make love to M"—"B has mother complex"—"Oedipus complex"—"B loves M"—"B desires incest")

Related comments:

- (1) Connotation of oedipal involvement on other cartoons.

(Examples: "B is very much in love with his M and stays near her all the time"—"By shaking M's collar which represents M, B is attempting to forestall a mother complex" — "B was very disappointed because M had not shown her love for him"—"He wants to come between his M and P")

- (2) Sexual or love reference involving B and M on other cartoons.

(Examples: "M seems to be in ecstasy while mothering B. Yet it is natural. However the drawing is unnecessarily suggestive."—"B is using his sex urges for the first time, he is trying to get the relationship of his body to hers"—"He probably wants M for other reasons than food" — "Much more of this (discovering sex) and M had better watch out") (One point per cartoon including comments (1) and (2)).

- (3) B competitive toward P (on spontaneous story for cartoon X).

(Examples: "B thinks he can do almost anything P does" — "B figures he'll be as big or better than P"—"B figures he got the best of P there" — "B wants to be greater than P")

Positive relationships

Resentment over oral deprivation
(Factor II-C)

Overt hostility toward sibling and mother (Factor VIII-A)

Avoidance (DPI-I)

Intellectualization (DPI-VI)

Youngest of three or more siblings

Verbal (SVI)

Poor eyesight

Negative relationships

(None)

Interpretation.

An individual scoring high on this factor exhibits the classical symptoms of an unresolved Oedipus complex—sexual desire for the mother coupled

with an intensely competitive attitude toward the father, who is viewed as selfish and punitive. His frustration takes the form of hostility toward the mother in berating her for deprivation of oral supplies and favoring older siblings. Choice of the avoidance DPI alternative on Cartoon I is probably more a response to the phrase "here he's making up for lost time" rather than the avoidance aspect itself ("he often forgets to come to eat"). A positive correlation with Verbal on the Strong suggests an interest pattern typical of advertising men, lawyers, and author-journalists.

*Factor IV-B. Disguised
Oedipal Involvement*

Inquiry:

- 3b B is most unhappy over the idea that M and P seem to be ignoring him on purpose.
- 4d B suspects M and P are purposely depriving him of attention.
- 5 P will do nothing.
- 7 (Answer to question "why?" in conjunction with choice of alternative a): Father-son relationship is healthy.

(Examples: "It would show B and P are getting along so well"—"Because father and son should be very close companions"—"Two males are more common"—"He shows too much attachment to M") (Answer to question "why?" in conjunction with choice of alternative b): B and M belong together more naturally.

(Examples: "It would be more natural because B is a boy and M is a girl"—"It would seem more logical for the pup to be with his M"—"Maternal love is considered stronger than paternal love"—"It's the most usual picture")

Preference:

Like

Theme (including item 1 of inquiry):

B wants attention (*not* affection).

(Examples: "B is no longer parents' only interest"—"B wants attention"—"B is spoiled") (Note: Attention theme must be devoid of oedipal connotation to be scored.)

Positive relationships

- Minimizing castration anxiety (Factor VI-B)
- Projection (DPI-VII)
- Mother born in United States
- Dislike for oedipal jokes (Humor)
- Recall of oedipal cartoons (Humor)
- Sibling hostility (Social Perception)

Negative relationships

- Overt hostility toward sibling and mother (Factor VIII-A)
- Reaction formation (DPI-I)
- Intellectualization (DPI-VII)
- Intellectual orientation (Social Perception)

Interpretation.

The oedipal involvement here may be just as strong but the emphasis is on its defensive distortion in the guise of mere desire for parental attention. The defensive quality is apparent in several connections: a mildly paranoid, projecting attitude; rationalization of the longing for maternal closeness; minimizing of castration fears which normally accompany oedipal conflict; and indirect rather than direct manifestation of sibling rivalry. An underlying concern over oedipal problems is also reflected in the humor data, where content in that category is disliked yet not easily forgotten.

MASTURBATION GUILT (Cartoon V)

*Factor V-A. Fear of Punishment
for Masturbation*

Inquiry:

- 2b When B is older he will be enjoying himself but a little worried about this situation.
- 4 Yes, implication of yes, possibly (B does fear something might happen to him).

- 1 (In answer to question "what?" provided first part of item is answered "yes"): B is afraid of being discovered.
- 5 M will scold, warn, or stop him.
- 6 P will scold, warn, or stop him.

Theme:

B is afraid of being discovered.

(Examples: "One eye peeks out to watch for parents"—"B is watching to see if anybody is looking at him")

Positive relationships

Rejection in favor of sibling (Factor VIII-C)

Qualification of pervasive guilt (Factor IX-C)

Negative perception of self and father (Factor X-B)

Projection (DPI-II)

Intellectualization (DPI-III)

Political (AVL)

Mother reported as scolding

Recall of oedipal cartoons

(Humor)

Hay fever

Negative relationships

Sugar-coating (Factor I-C)

Choosing obvious neutral responses (Factor III-B)

Avoidance (DPI-VI)

Business detail (SVI)

Recall of anal cartoons (Humor)

Interpretation.

The focus of this factor is anxiety in connection with masturbation. Fears of being discovered and punished for the act are salient. Parents, especially mother, are pictured as punitive. In general, self-abasement, guilt, and feelings of rejection are expressed without any attempt to avoid or cover up. One outlet is suggested by a positive correlation with Political on the Allport-Vernon-Lindzey — attainment of personal power and influence over others offers promise of reassurance in the face of one's own basic fearfulness.

*Factor V-B. Concern over
Sexual Maturation*

Inquiry:

Ic B feels mixed up and guilty.

3 B is thinking about T.

4 (In answer to question "what?" provided first part of item is answered "yes"): B doesn't know what he's afraid of.

5 M will explain, advise.

6 P will explain, advise, or praise B.

Theme:

B is growing up and thinking about girls.

(Examples: "B discovered something new about himself and the sensations are pleasant, undoubtedly will develop into a healthily sexed dog and start noticing females in a different light"—"B is maturing into an older dog and realizes certain organs of his body operate to arouse a pleasant sensation when he sees a female dog"—"This is the first stage, from there through other stages to a family of his own")

Positive relationships

Supply-seeking (Factor II-B)

Overt hostility toward sibling and mother (Factor VIII-A)

Major in pre-med

Theoretical (AVL)

Proud (Interpersonal Mechanisms)

Religious training in childhood

Negative relationships

Projection (DPI-X)

Interpretation.

Sexual maturation appears to be accompanied by feelings of guilt, confusion, and free-floating anxiety, possibly exacerbated by a strongly religious upbringing. Although there is evidence of interpersonal frustration and discontentment, both parents are presented as trying to be helpful and S is unwilling to ascribe blame to his family. Theoretical interests in science and the discovery of truth, along with choice of a medical career, may very well reflect the need to satisfy curiosity, initially associated with disturbing thoughts about differences between the sexes.

*Factor V-C. Denial of
Masturbation Guilt*

Inquiry:

- 1a B feels happy, without a care in the world.
- 2a When B is older in this situation he will feel happy, without a care in the world.
- 3 B is not thinking of anyone here.
- 5 M will say nothing.
- 6 P will say nothing.

Themes:

- (1) Denial of B's concern.

(Examples: "B is washing himself just as all dogs do. There is nothing funny or no sex involved"—"B chasing after fleas comes across his genitals and is licking them. The taste is probably salty and B likes salt; he has no other view in mind and is merely reacting as all dogs do to salt solution on their bodies"—"This behavior is not intrinsically immoral"—"B sees nothing wrong in it")

- (2) B chasing fleas, cleaning himself.

(Examples: "B is looking for fleas, trying to get rid of fleas by licking them"—"Shows he's been brought up well by family, taught to keep himself clean") (Note: not "cleaning" and sexual reference mixed together in same story)

Positive relationships

- Choosing obvious neutral responses (Factor III-B)
- Minimizing castration anxiety (Factor VI-B)
- Reaction formation to sibling rivalry (Factor VIII-B)
- Intellectualization (DPI-XI)
- Business contact (SVI)
- Father reported as lecturing

Negative relationships

- Exploitation (Factor III-A)
- Father as preferred identification object (Factor VII-A)
- Rejection in favor of sibling (Factor VIII-C)
- Thumbsucking

Interpretation.

Here again is a constellation which reflects defensive denial of an underlying conflict. Guilt over masturbation is handled by evasion, negation, and adoption of a positively-toned approach. The concomitant relationships bear out the preference for avoidant over expressive items, especially on the anal sadism, castration, and sibling rivalry dimensions. Interest in business contact on the Strong may indicate that the position of salesman, emphasizing extraverted activity, helps to divert attention from guilt and other self concerns.

CASTRATION ANXIETY (Cartoon VI)

*Factor VI-A. Overwhelming
Castration Conflict*

Inquiry:

- 1a B feels terrified that he's going to be next.
- 3b B is thinking desperately about a way to save his own tail.
- 3c B thinks he might look better if his own tail is cut off.
- 4—Yes ("why?"): B would rather have his own tail cut off than go through the suspense of wondering if it will happen to him.
- 4—No ("why?"): B is too frightened to have it cut off right away.

(Examples: "No, because he's afraid to get his tail cut off"—"No, because B doesn't know if it would hurt or not"—"No, he will be afraid of the pain seen in T"—"No, because he does not want to be afraid")

- 5 P most likely arranged for T's tail to be cut off.
- 6b Other dogs in the neighborhood will make fun of T's short tail.

Preference:

Worst

Themes:

- (1) B is wondering if he's next; doesn't know he's next.

(Examples: "He is a little worried about his meeting the same fate, and

also feels sorry for T"—"B wonders when it will happen to him"—"B is somewhat disgusted but more afraid and worried, for his turn is next")

(2) B feels sorry for T.

(Examples: "B is very sad because T is going to have her tail cut off"—"He thinks, 'Poor Tippy'" — "He knows that it is being done for her beauty and can't do anything about it but he is sympathizing with her")

Positive relationships

Sugar-coating (Factor I-C)

Projection (DPI-V)

Intellectualization (DPI-VI)

Major in physical or natural science

Intellectual orientation (Social Perception)

Production manager (SVI)

Dislike for oral sadistic jokes (Humor)

Negative relationships

Major in liberal arts

Interpretation.

A high-scoring individual on this factor is caught in the throes of an intense psychic struggle over his masculinity as well as fears of physical injury. He dreads symbolic castration to the point of considering assumption of a passive, effeminate role, yet such a solution is itself too fraught with anxiety to be acceptable. Defensive reactions typically involve an intellectualized approach: dependency needs are glossed over; frightening experiences are justified as building up resistance to future tribulations; and standards of intellectual achievement are stressed. Only in connection with masturbation guilt, presumably a contributing factor to the castration anxiety, is there recourse to a more primitive defense, namely projection. The choice of a physical or natural science as field of specialization seems consistent with the detached, intellectualized approach. A vocational interest pattern corresponding to production manager may reflect another kind of outlet —

achieving control and power over others as reassurance against one's own fears of masculine inadequacy.

*Factor VI-B. Minimizing
Castration Anxiety*

Inquiry:

1c B feels curious but calm.

2c B suspects T is being improved in some way.

3a B is not particularly worried about his own tail.

4—No ("why?"): B is not concerned about his own tail.

(Examples: "No, he doesn't expect it himself"—"I don't think he thinks anything about his own tail"—"No, I don't think he is worried"—"Doesn't concern him, not old enough to think about it")

Themes:

(1) B is hanging around and watching curiously.

(Examples: "B saw T in a strange situation and stood around to see the results"—"In general, he will be a curious bystander" — "B seems naively interested in this procedure; a smarter dog would not be in the vicinity")

(2) B minimizes the whole situation.

(Examples: "As T is blindfolded, B watches T lose her tail with apparently no emotion"—"T's going to get hurt but B is not at all worried; he doesn't even care")

(3) T afraid, puzzled, upset.

(Examples: "B is not apprehensive because nothing has yet happened, but T is because she is blindfolded"—"Poor T is terrified and will surely let out a howl when he loses his tail"—"T is nervous since this never happened to her before"—"T is scared")

Positive relationships

Attempted denial of anal preoccupation (Factor II-C)

Disguised oedipal involvement (Factor IV-B)

Denial of masturbation guilt (Factor V-C)

Intellectualization (DPI-I)
 Reaction formation (DPI-VI)
 Sibling hostility (Social Perception)
 Raised in large city
 Punitive (Interpersonal Mechanisms)
 Respectful (Interpersonal Mechanisms)

Negative relationships

Regression (DPI-I)
 Projection (DPI-V)
 Intellectualization (DPI-VI)
 Production manager (SVI)
 Hay fever
 Back pain

Interpretation.

This is another in the series of defensive factors, several of which are correlated positively with it. Negative relationships with a number of expressive DPI mechanisms go along with the interpretation that the items indicate desire to minimize involvement over castration. There is also evidence that the avoidant quality may on occasion turn into reaction formation, emphasizing possible advantages of relinquishing masculinity. The interpersonal data, revealing punitive, derogatory attitudes toward others mingled with deferential respectfulness, likewise fit the view that conflict over playing a dominant masculine role underlies the factor responses.

IDENTIFICATION PROCESS (Cartoon VII)

Factor VII-A. Father as Preferred Identification Object

Inquiry:

- 1 P talks like that to B. ("M and P" not scored on any of these items)
- 2 B is most likely to obey P.
- 3 B is imitating P.
- 4 B would rather pattern himself after P.
- 5 B's disposition, actually, is most like P's.
- 6c B would get mad and sulk if he

were in the position of the toy dog.

Preference:

Like

Themes:

(1) B is imitating P.

(Examples: "B has at last found someone to boss around. He is trying to act like his father"—"B is pretending he's a papa dog and telling some smaller dog man to man stuff"—"B is feeling fatherly toward the toy dog, playing house")

(2) B is jealous of the toy, dislikes it.

(Examples: "The vicious feeling that he is seeing a new aspirant for attention in his neighborhood makes him mad"—"B looks angry at the toy, as if he dislikes it"—"B thinks this dog is going to join his family, he doesn't want any more dogs in the family and he is going to growl at the dog and hit it to scare it away")

Positive relationships

Guilt-ridden hostility toward sibling (Factor IX-B)
 Avoidance (DPI-VII)
 Projection (DPI-X)
 Total number of siblings
 Loving (Interpersonal Mechanisms)
 Mother reported as scolding
 Thumbsucking
 Smoking

Negative relationships

Choosing obvious neutral responses (Factor III-B)
 Denial of masturbation guilt (Factor V-C)
 Regression (DPI-II)
 Mother reported as using deprivation for punishment
 Business detail (SVI)
 Raised in large city

Interpretation.

The focus here is on father as the preferred identification object. Father is seen as the decisive disciplinarian, despite the fact that mother frequently scolds. However, attempts to pat-

tern after the father are not devoid of anxiety, for the role carries with it a disturbing quality of aggressiveness. Manifestations of this disturbance are guilt over hostility toward siblings (typically in a large family); choice of the avoidance response to aggression in the DPI; and possibly a need to appear loving in interpersonal situations. Other indications that the identification process is not wholly gratifying are the projection of feelings of inadequacy onto the family, and the reported thumbsucking in childhood.

Factor VII-B. Mother as Preferred Identification Object

Inquiry:

- 1 M talks like that to B. ("M and P" not scored on any of these items)
- 2 B is most likely to obey M.
- 3 B is imitating M.
- 4 B would rather pattern himself after M.
- 5 B's disposition, actually, is most like M's.

Theme:

B is imitating M.

(Examples: "B is doing the same thing to the toy dog as his mother does to him"—"Here B is playing like M and pretending that the toy dog is himself")

Positive relationships

Overt hostility toward sibling and mother (Factor VIII-A)

Rejection in favor of sibling (Factor VIII-C)

Negative perception of self and father (Factor X-B)

Narcissism (Factor XI-C)

Negative relationships

Reaction formation to sibling rivalry (Factor VIII-B)

Guilt-ridden hostility toward sibling (Factor IX-B)

Avoidance (DPI-I)

Avoidance (DPI-IX)

Loving (Interpersonal Mechanisms)

Social service (SVI)

Recall of anal cartoons (Humor)

Interpretation.

In this factor mother is the decisive figure and serves as preferred identification object. Disappointment and hostility toward her, especially in the context of her treatment of siblings, accompany the feminine identification. Satisfactory heterosexual relationships are unattainable and instead a purely narcissistic approach to the opposite sex is pursued. In general, the setting is an unhappy one, with feelings of worthlessness and deprecation of father as a model both prevalent. There is little capacity for love or interest in gratifying the needs of others.

Factor VII-C. Evasion of Identification Issue

Inquiry:

- 1 T talks like that to B.
- 2 B is most likely to obey T.
- 3 B is imitating T.
- 4 B would rather pattern himself after T.
- 5 B's disposition, actually, is most like T's.
- 6d B would have an impulse to start fighting if he were in the position of the toy dog.

Theme:

B is playing, pretending.

(Examples: "B is playing with the toy dog, it looks as though he's having fun"—"B again is playing in the land of make-believe. Here he thinks he is company commander giving orders to his men")

Positive relationships

Choosing obvious neutral responses (Factor III-B)

Avoidance (DPI-III)

Motion sickness

Political (AVL)

Negative relationships

Negative perception of self and father (Factor X-B)

Intellectualization (DPI-II)

President of manufacturing company (SVI)

Interpretation.

One plausible view of this constellation is that S, rather than seeking to identify with a sibling figure, chooses the "Tippy" alternatives as a way out of selecting either "Mama" or "Papa." The spontaneous story, typically evoking thoughts about aggression and parents, is glossed over as innocuous play or make-believe. By the end of the inquiry an aggressive breakthrough does occur, however. The correlated evasive and avoidant reactions to the anal picture (also involving aggression and parents), together with a high value attached to political power, all support a defensive interpretation.

SIBLING RIVALRY (*Cartoon VIII*)*Factor VIII-A. Overt Hostility toward Sibling and Mother*

Inquiry:

- 1d B wants to run away to spite M and P.
- 2c B feels T actually deserves to be punished instead of praised.
- 3a B feels M is paying more attention to T.
- 4c B sees this very often.
- 6 B is most angry at M.

Theme:

B has hostile feelings.

(Examples: "B doesn't like T because she's a 'Parent Dog'; he'll get even"—"B hates T now"—"He gets madder and madder as he watches the proceedings")

Related comment:

Cartoon IV, Item 4a; B suspects M and P are planning an addition to the family.

Positive relationships

- Oral craving (Factor I-A)
- Supply-seeking (Factor II-B)
- Resentment over oral deprivation (Factor II-C)
- Undisguised oedipal involvement (Factor IV-A)
- Concern over sexual maturation (Factor V-B)

Mother as preferred identification object (Factor VII-B)

Negative perception of self and father (Factor X-B)

Projection (DPI-III)

Regression (DPI-IV)

Regression (DPI-VIII)

Reaction formation (DPI-X)

Punitive (Interpersonal Mechanisms)

Poor eyesight

Recurrent bedwetting

Fainting

Total number of physical complaints

Negative relationships

Playfulness (Factor II-A)

Choosing obvious neutral responses (Factor III-B)

Disguised oedipal involvement (Factor IV-B)

Avoidance (DPI-II)

Reaction formation (DPI-VIII)

Reaction formation (DPI-XI)

Social service (SVI)

Interpretation.

Direct expression of hostility toward sibling and parents, especially mother, characterizes this factor. Regressive reactions to such hostile impulses appear in the DPI and the reported childhood physical complaints of recurrent bedwetting and fainting. Most striking is the result that scores on *Overt Hostility toward Sibling and Mother* are correlated positively with scores on a host of other factors expressive of disturbance in a variety of areas, and negatively with several defensive ones. It seems almost as though this factor can serve as a general index of undefended conflict, which is not surprising since attitudes toward the family (tapped by *Cartoon VIII*) are a product of many specific interactions (tapped by other pictures). Even the most general indicator of possible psychosomatic involvement, total number of physical complaints, is correlated significantly in the positive direction.

*Factor VIII-B. Reaction Formation
to Sibling Rivalry***Inquiry:**

- 1b B wants to bark happily at the group and join them.
- 2a B feels that actually T fully deserves the praise.
- 3c B feels that both parents are paying the same amount of attention to T.
- 5b B thinks M and P really love him about the same as they do T.
- 6 B is not angry at anyone.

Preferences:

Like { (one point for each)
Best {

Themes:

(1) Emphasis on B's happiness.
(Examples: "B looks like he is happy. In a minute he will probably join them so he can get into the picture too"—"T is being rewarded for some commendable act; B is looking on with admiration"—"He is proud of his family. They have shortcomings but so what? He wishes he had a camera so that he could record this tender scene for posterity")

- (2) Denial or minimizing of B's concern.

(Examples: "So they're trying to make me jealous, eh? Well, I don't give a damn and they can't make me jealous anyway"—"B is now becoming more used to the idea of not being the center of attraction and doesn't get angry so easily"—"B is not disturbed but simply wondering why T is getting so much attention"—"B is watching with apparently little or no feeling while M and P caress T")

Positive relationships

- Playfulness (Factor II-A)
- Choosing obvious neutral responses (Factor III-B)
- Denial of masturbation guilt (Factor V-C)
- Reaction formation (DPI-VIII)
- Reaction formation (DPI-XI)
- Raised in large city

- Mother perceived as final authority in discipline
- Mother reported as using corporal punishment
- Headaches
- Interest maturity (SVI)

Negative relationships

- Fear of punishment for masturbation (Factor V-A)
- Mother as preferred identification object (Factor VII-B)
- Negative perception of self and father (Factor X-B)
- Regression (DPI-I)
- Regression (DPI-VIII)
- Regression (DPI-XI)
- Projection (DPI-VIII)
- Father perceived as final authority in discipline
- Excessive drinking

Interpretation.

The salient feature here is the effusive effort to dispel uncomfortable hostility by proclaiming that everything is wonderful—one big, happy family with B not angry or concerned. Consistent with this avoidant attitude are the positive relationships with defensive factors from several other cartoons. There is a hint that the reaction may stem from the necessity to placate a highly dominant, punitive mother. Subduing hostile thoughts apparently takes its toll, however, in the form of frequent headaches.

*Factor VIII-C. Rejection in Favor
of Sibling***Inquiry:**

- 1d B wants to run away to spite M and P.
- 2b B feels that T actually deserves some praise but not that much.
- 4b B sees this fairly often.
- 5c B thinks M and P really love him less than they do T.

Preference:

Worst

Theme:

Strong feelings of rejection.
(Examples: "B seems to be the

black sheep of the family. He is very unhappy at being left out of things"—"B will either break up completely and start weeping and then get mad or he'll become a delinquent dog"—"B probably thinks he is not loved by anyone and is contemplating running away at this very moment"—"B is hurt, angry, and has a feeling of betrayal"—"He feels left out and unwanted, nobody cares about him and they don't even see him")

Positive relationships

Fear of punishment for masturbation (Factor V-A)
 Mother as preferred identification object (Factor VII-B)
 Negative perception of self and father (Factor X-B)
 Intellectualization (DPI-II)
 Projection (DPI-VIII)
 Regression (DPI-VIII)
 Avoidance (DPI-XI)
 Sole sibling is younger brother
 Father reported as scolding
 Excessive smoking
 Nailbiting

Negative relationships

Denial of masturbation guilt (Factor V-C)
 Avoidance (DPI-VIII)
 Reaction formation (DPI-VIII)
 Reaction formation (DPI-XI)
 Regression (DPI-X)
 Projection (DPI-XI)
 Mother reported as using corporal punishment
 Supportive (Interpersonal Mechanisms)
 Social service (SVI)
 Raised in large city

Interpretation.

A deep-seated feeling of rejection pervades this factor. Parents are viewed as favoring the younger brother to such an extent that S has fantasies of running away to retaliate. Preferred defense mechanisms in the sibling rivalry conflict are typically primitive, i.e., projection and regression. A scolding father is perceived as an inadequate ego ideal figure and

instead mother is chosen as the preferred identification object. Expressed fear of punishment for wrongdoing, plus the tension indices of excessive smoking and nailbiting, round out the picture of despondency.

GUILT FEELINGS (Cartoon IX)

Factor IX-A. Partial Denial of Guilt Inquiry:

- 2c B is hardly bothered at all by his conscience, just afraid of what will be done to him.
 4c The situation couldn't be helped.
 5c B hardly feels guilty at all.
 6 (1) B will be morose, . sad, sulky.
 (Examples: "B will cry and give up hope"—"B will feel very badly"—"He'll worry a lot but not do much"—"B will go off and sulk"—"He'll run to his M and cry on her shoulder")
 (2) B might die.
 (Examples: "He'll crawl in a hole and die!"—"B might kill himself")
 7c B will feel bad for a little while and then go out to play.

Theme:

B in relation to the hereafter or God; B's own death.

(Examples: "B shakes and shudders as he thinks of his horrible past as he can see an angel condemning him to Hell. God have mercy on B"—"B's afraid that his affair with T will keep him from a dog's heaven"—"If he kills himself he certainly will go to a bad place—Hell!"—"B has just discovered religion and has come to wonder and worry about death")

Positive relationships

Avoidance (DPI-III)
 Reaction formation (DPI-IV)
 Regression (DPI-VIII)
 Major in liberal arts
 Raised in large city

Negative relationships

Projection (DPI-IV)

Regression (DPI-V)
Regression (DPI-VI)

Interpretation.

Here a striking split is observed between spontaneous, open-ended responses and the highly structured, multiple-choice inquiry answers. On the former there is evidence of a very depressive concern over death and its aftermath; on the latter guilt is steadily denied. The pattern of defense preferences is similarly inconsistent: avoidant alternatives tend to be selected and expressive ones rejected, yet regression suddenly crops up in connection with sibling rivalry. In general the factor can be seen as an unsuccessful attempt to deny deeply rooted moral anxieties. It appears to occur most typically in students from large cities who become liberal arts majors in college.

Factor IX-B. Guilt-Ridden Hostility toward Sibling.

Inquiry:

- 1 Hostility toward T.
(Examples: "He might have bitten T"—"He beat T up"—"B has murdered T"—"B did something to be blamed on T")
- 2a B's conscience is so strong he's practically paralyzed.
- 3 The actions of the pointing figure remind B of T.
- 4a B himself is really to blame for his feeling this way.
- 5a B feels very guilty.
- 6 B will wait for the worst; expect punishment.
(Examples: "B will sit back and let himself be blamed for his crime" — "He'll pay for his wrong"—"He'll go home and face the music"—"He'll take it like a man")
- 7a B will have this feeling as long as he lives.

Theme:

Hostility toward T.

(Examples: "I would say that B had murdered T"—"He thought it

over and decided to kill T"—"He has been extremely cruel to T and now his conscience bothers him"—"B just stole T's food and was congratulating himself on his exploits")

Positive relationships

Father as preferred identification object (Factor VII-A)
Regression (DPI-III)
Regression (DPI-VIII)
Projection (DPI-IV)
Speech difficulties

Negative relationships

Mother as preferred identification object (Factor VII-B)
Avoidance (DPI-III)
Mother reported as using corporal punishment
Major in liberal arts

Interpretation.

The extremely intense guilt reaction reflected in this factor is tied specifically to violent, hostile thoughts about a sibling figure. The guilt is so strong and lasting that a need to be punished (presumably as expiation) manifests itself. The aggressive fantasies are compatible with preferred father identification on Cartoon VII, but their disturbing quality shows up in other ways besides guilt. The presence of speech difficulties, linked to blocking of hostile expression as mentioned earlier, and regressive defense preferences on the anal sadism and sibling rivalry dimensions betray a serious conflict.

Factor IX-C. Qualification of Pervasive Guilt

Inquiry:

- 2b B's conscience is bothering him somewhat, but he's mostly afraid of what will be done to him.
- 5b B feels fairly guilty.
- 7b B will feel bad every now and then.

Preference:

Worst

Theme:

Guilt will be long-lasting; depressed, despondent.

(Examples: "He comes to the conclusion that he is a very worthless character compared to other dogs"—"He feels very low"—"Here B is beaten and dejected. He feels that no one loves him and he might as well leave"—"He can't sleep. It keeps coming back to torment him. He did something he shouldn't of and now he keeps thinking of it")

Related comments:

Reference to Cartoon IX or conscience.

(Examples: "He wants to be the envy of all the other dogs, also compensate for his wrongdoing"—"The trouble with his conscience is over now and he has forgotten it"—"B feels guilty that he's not barking or in some way warning T"—"Nevertheless B is a neat, well-mannered dog and is putting dirt around it. He is ashamed of what he did") (one point per cartoon)

Positive relationships

- Exploitation (Factor III-A)
- Fear of punishment for masturbation (Factor V-A)
- Avoidance (DPI-I)
- Regression (DPI-V)
- Regression (DPI-VII)
- Projection (DPI-X)
- Father reported as using corporal punishment
- Major in pre-med
- Major in social science
- Extracurricular activities

Negative relationships

- Choosing obvious neutral responses (Factor III-B)
- Reaction formation (DPI-IV)
- Intellectualization (DPI-V)
- Avoidance (DPI-XI)
- Father reported as using deprivation for punishment
- Mother born in United States
- Sole sibling is an older brother

Interpretation.

The guilt here is also strong and pervasive but the defensive reaction is one of qualification rather than denial or avoidance. Guilt feelings

are acknowledged on the inquiry to be troublesome but not overly so (contradicted by evidence from themes, preferences, and related comments). The family situation appears particularly difficult. Father is fear-inspiring and physically punitive; mother, foreign-born, has no offsetting supports to provide on her part. Hostility toward the parents is indicated in the positive correlation with *Exploitation*, and in the DPI feelings of inadequacy are projected onto the family. Fear of punishment for masturbation, regressively defended against, poses a real problem and may have helped to direct interest into the medical area. Social science may have an appeal because of the opportunity to work off guilt feelings. The prevalence of extracurricular activities suggests a turning away from the family for interpersonal gratifications.

EGO IDEAL (Cartoon X)

Factor X-A. Overtly Positive Perception of Self and Father

Inquiry:

- 1 The dream figure reminds B of P.
- 2 P stacks up favorably against the dream figure when B compares them.
- 4a B probably feels he has a very good chance to grow up to be like the dream figure.
- 5a Actually B's chances of growing up to be like the dream figure are very good.

Preferences:

- | | |
|------|------------------------|
| Like | } (One point for each) |
| Best | |

Theme:

P esteemed positively.

(Examples. "B has always admired the gracefulness of his father. He dreams he will grow up to be as fine a dog as P and maybe a little better"—"B is dreaming that he is grown up and a famous hunting dog, strong, proud, and smart. Perhaps his

father was once a dog like this or he just wants to be this of his own volition"—"B dreams that someday he will be the great hunter that his father is")

Positive relationships

- Sugar-coating (Factor I-C)
- Mother-surrogate as love object (Factor XI-A)
- Reaction formation (DPI-IV)
- Interest maturity (SVI)

Negative relationships

- Exploitation (Factor III-A)
- Reaction formation (DPI-III)
- Youngest of 3 or more siblings
- Major in business, or pre-law
- Father reported as scolding

Interpretation.

At first glance these factor variables are capable of two opposing interpretations: a high-scoring S really patterns himself successfully after a desirable father figure, or else he is expressing a set of wish-fulfilling fantasies which have little or no actual basis. The relationships of the factor to criterion variables clarify the dilemma. Positive correlations with *Sugar-Coating* on oral eroticism and reaction formation on the oedipal cartoon ("he'll cheer up when he realizes that M and P love each other so much") both point to the defensive interpretation. Furthermore, the oedipal-based choice of mother as love object on Cartoon XI contra-indicates a wholesome resolution of the identification process. The suggestion of stable vocational interests over time does not add much, since it may reflect either true maturity of interests or some form of test-taking defensiveness.

Factor X-B. Negative Perception of Self and Father

Inquiry:

- 2 P is not as good as the dream figure when B compares them.
- 3c B's main reason for wanting to be like the dream figure is that then he would be loved more by M and P.

5c Actually B's chances of growing up to be like the dream figure are very poor.

6a B has this kind of dream very often.

Theme:

B indulging in wish-fulfilling fantasy.

(Examples: "B dreams he is a super dog, very handsome, strong, and perfect form"—"By golly, I'm going to be the best bird dog possible"—"Here he is a knight in shining armor, all-powerful"—"He's dreaming of what he'd like to be when he grows up")

Related comments:

Any reference on Cartoon XI back to Cartoon X.

Positive relationships

- Oral craving (Factor I-A)
- Oral rejection (Factor I-B)
- Supply-seeking (Factor II-B)
- Exploitation (Factor III-A)
- Fear of punishment for masturbation (Factor V-A)
- Mother as preferred identification object (Factor VII-B)
- Overt hostility toward sibling and mother (Factor VIII-A)
- Rejection in favor of sibling (Factor VIII-C)
- Narcissism (Factor XI-C)
- Sole sibling is a younger sister
- Youngest of 3 or more siblings
- Major in business or pre-law
- Father reported as scolding
- Total number of physical complaints

Negative relationships

- Playfulness (Factor II-A)
- Choosing obvious neutral responses (Factor III-B)
- Evasion of identification issue (Factor VII-C)
- Reaction formation to sibling rivalry (Factor VIII-B)
- Heterosexual fantasy (Factor XI-B)
- Avoidance (DPI-II)
- High grade-point average

Interpretation.

In this factor both father and self

are downgraded and life is viewed pessimistically. The only recourse in an otherwise abject mental state is regressive fantasy, where dreams of love and glory are freely indulged. However, the fanciful flights cannot effectively cushion the harsh impacts of reality, judging by the large number of positive relationships with factors expressive of disturbance. The pattern also holds for two very different types of sibling constellation. Like VIII-A this factor can almost serve as a general index of undefended conflict. Even total number of physical complaints is related as before. Negative correlations again occur with evasive, avoidant factors. One of the unique features here, not present previously in VIII-A, is the toll apparently taken in academic performance—high scorers tending to have poorer grades.

LOVE OBJECT (Cartoon XI)

Factor XI-A. Mother-Surrogate as Love Object

Inquiry:

- 1 B is dreaming about M.
- 2 The dream figure reminds B of M.
- 3b B is most attracted by the possibility that the dream figure looks like M, which reminds him of the good old days.
- 4b B is so contented because he feels M will comfort him.
- 5 In B's mind M is comparable to the dream figure.
(Examples: "M is pretty close to the dream figure"—"M compares fairly well"—"They're similar"—"Okay"—"M is not as beautiful but just lovable in a different way")

Theme:

Dream figure is M or resembles M.

Positive relationships

- Resentment over oral deprivation (Factor II-C)
- Evasion of identification issue (Factor VII-C)

- Overtly positive perception of self and father (Factor X-A)
- Recall of oedipal cartoons (Humor)
- Excessive drinking
- Childhood discipline reported as strict
- Social science major

Negative relationships

- Playfulness (Factor II-A)
- Intellectualization (DPI IV)
- Reaction formation (DPI-VII)

Interpretation.

The choice of a motherly love object here is accompanied by infantile longings for care and comfort. Though mother is portrayed favorably, there are several signs that actually all was not well in early childhood. Resentment over oral deprivation, going along with the current admission of excessive alcohol consumption, and reported strictness of discipline together suggest that the maternal fixation grew out of frustration rather than overindulgence. Oedipal and identification problems very likely also contributed to the desire for a mother surrogate. The salience of oedipal humor, plus defensiveness in connection with identification and ego ideal figures, tend to substantiate such a view.

Factor XI-B. Heterosexual Fantasy

Inquiry:

- 2 The dream figure doesn't remind B of anyone.
- 3c B is most attracted by the possibility that the dream figure looks like someone else, whom he would make happy by giving her all his love.
- 4c B is so contented because he feels the dream figure will be delighted by his attentions.
- 5 In B's mind there is no similarity between M and the dream figure.
(Examples: "Not very alike"—"No similarity between them"—"Not at all"—"Out of the question")

- 6 No, B doesn't want to be like the dream figure.
- 6 (answer to question "why?"): B doesn't like effeminate boys.
(Examples: "No, he's no sissy" — "No, he's not really a fairy" — "To be like her wouldn't be masculine")

Preference:

Best

Theme:

B thinks the dream figure is unattainable.

(Examples: "He feels that she is perfection and he has little chance of attaining her" — "In real life he knows he can never marry anything as nice as this dog" — "He likes her very much, hopes she likes him, is afraid she doesn't")

Positive relationships

- Reaction formation (DPI-III)
- Reaction formation (DPI-IV)
- Reaction formation (DPI-VIII)
- Reaction formation (DPI-XI)
- Avoidance (DPI-VIII)
- Projection (DPI-XI)

Negative relationships

- Negative perception of self and father (Factor X-B)
- Regression (DPI-IV)
- Regression (DPI-VIII)
- Regression (DPI-XI)
- Projection (DPI-III)
- Projection (DPI-VIII)
- Reaction formation (DPI-X)
- Recall of oedipal cartoons (Humor)
- Large amount of punishment in childhood reported

Interpretation.

This factor expresses heterosexual interests but they are confined to the level of vague fantasy. S has no definite love object in mind during the highly pleasant fantasies, and mother is specifically excluded. The unrealistic quality is confirmed by both defenses on Cartoon XI which correlate positively: the projection alternative

states that B isn't concerned about not getting a dog like his dream figure because beautiful females are sometimes inadequate; and the reaction formation has B merely dreaming of adding another prospect to a long list of female conquests. S also seems determined to portray the family group in a rosy light, with his repeated preferences for reaction formation on parental dimensions. The inability to recall oedipal humor or to say bad things about self and father are all consistent with this tendency.

Factor XI-C. Narcissism

Inquiry:

- 1 B is dreaming about himself.
- 2 The dream figure reminds B of himself.
- 3a B is most attracted by the possibility that the dream figure looks like himself, which would increase his pride.
- 4a B is so contented because he feels everyone will admire him.
- 6 Yes, B would rather be like the figure in his dream.
- 6 (answer to question "why?"): Because the dream figure is a superior kind of dog.
(Examples: "Yes, he could show up everyone else" — "In some ways perhaps because she is gentle and sympathetic" — "Yes, because it's good-looking")

Theme:

The dream figure is B or resembles B.

Positive relationships

- Oral rejection (Factor II-B)
- Mother as preferred identification object (Factor VII-B)
- Negative perception of self and father (Factor X-B)
- Regression (DPI-IV)
- Regression (DPI-VIII)
- Regression (DPI-XI)
- Projection (DPI-III)
- Projection (DPI-VIII)

Middle of three or more siblings
 Total number of siblings
 Sleep difficulties
 Recurrent bedwetting
 High grade-point average

Negative relationships

Reaction formation (DPI-IV)
 Reaction formation (DPI-VIII)
 Reaction formation (DPI-XI)
 Avoidance (DPI-VIII)

Interpretation.

The individual who scores highly here has a narcissistic rather than heterosexual orientation. He is involved with pride over his own appearance, seeks to be admired by others, and expresses a willingness to relinquish the masculine role in order to achieve narcissistic gratification. A pessimistic attitude toward his own future indicates that these self-centered ambitions are not readily attainable. Parental influences predisposing toward a narcissistic, homosexual outlook are a rejecting mother with whom S identifies, and an inadequate father. The disturbed nature of the syndrome, which appears more likely to be found in large families, is revealed by the regressive physical complaints as well as defense preferences. The positive correlation with academic success probably means that, for the college student, good grades are a convenient means of boosting self-esteem and earning respect from others.

DISCUSSION

Any reader who has ploughed through the foregoing array of 30 factors is owed a few additional words of justification to make the venture seem worthwhile in retrospect. The proliferation of factors, while cumbersome to contemplate in its entirety, can be rationalized on two counts: 1) that's the way Blacky bounces, i.e., no broad cross-picture groupings emerged from the analyses; and 2) the test thereby makes available a vast amount of diversified in-

formation about the person who takes it.⁹

Actually there is a general observation pertinent to the series of factors—a division into expressive versus defensive. On each picture at least one factor reflects a defensive style of responding to the test. At times the style takes the form of denying or minimizing concern, at others deliberately selecting innocuous items, or overstating the merits of the situation, and so on. Often these defensive factors do intercorrelate positively (e.g., II-A, III-B, V-C, and VIII-B).

The interpretations represent the author's own best syntheses of the currently available data. They are subject to modification by others and, more important, by subsequent empirical evidence. However, it should be pointed out that the results frequently display a striking degree of coherence and consistency. The meaningful patterning typically made the identification of a factor easy and permitted ready consolidation of the positive and negative criterion correlations.

Advantages of the new scoring system over its predecessors are several (besides rendering obsolete perhaps a hundred previous studies!). It was achieved by means of an extended statistical approach which systematically explored all possible responses to the test stimuli. Any theme, open-ended answer, or related comment which does not appear somewhere as a factor variable either did not load on any factor or else was dropped from consideration because it occurred too infrequently (in 2% of the

⁹ This is not to say, of course, that the profile of factor scores for a given individual cannot be integrated into a valid clinical account of his personality dynamics. The absence of cross-picture groupings only implies that it is not possible to single out particular clusters which hold for large numbers of people. Readers interested in the clinical use of the Blacky Pictures are referred to the Manual of Instructions (Blum, 1950).

sample or less). The remaining variables can all be scored in relatively objective fashion. The numerical factor scores, obtained by simply adding the points from the separate variables loading on a factor, are convenient for use in correlational studies and allow a greater range than the former threefold categorization of disturbance into strong, moderate, weak or absent.

The factors themselves, besides conveying psychological and statistical entities, also help to differentiate some of the gross lumping of responses done earlier. For example, the previous scoring system treated oral craving and rejection content on Cartoon I as similarly indicative of disturbance. It now turns out that the two kinds of responses are contained in factors (I-A and I-B) orthogonal to one another. On the other hand, certain expulsive and retentive responses on Cartoon III, kept apart before, legitimately belong together under the rubric of *Exploitation* since they are shown to share an aggressive component. But perhaps the most notable advance is exposure of the factors to a great many diverse criterion data, all contributing to the establishment of their construct validity. Thus, a factor which on the surface appears capable of opposing interpretations is inevitably clarified in the process of examining its criterion relationships (see VII-C, X-A).

The reader should be reminded that the new scoring system is based upon responses of "normal" male undergraduates tested in group sessions.¹⁰ Extension to females, differ-

¹⁰ Group testing, though not always feasible practically, is preferred for purposes of research. In addition to economy of time, the group method has the advantage of minimizing differential interpersonal contact between examiner and subjects. The use of slides to project pictures and inquiry items on a screen also insures more uniform procedure by keeping constant the amount of time each S has to respond. Too, slide presentation of the inquiry allows the

ent age groups, socioeconomic levels, pathological populations, etc. all remain tasks for the future and other investigators. It would be very illuminating, for example, to compare factor scores among various clinical syndromes. Not only might further light be shed on the psychosexual dynamics of many types of mental illness, but also the results could be employed to assess the potential of these Blacky scores as indices both of personal adjustment (e.g., mental health in relation to high scores on expressive or defensive factors) and therapy prognosis. As stated in the Manual, the technique itself is primarily intended to reveal an individual's *predispositions* to psychosexual disturbance along a series of dimensions inherent in personality development, not to diagnose psychosis or classify into psychiatric categories. The assumption is made that everyone has to come to grips, at least to some extent, with the problems tapped by the Blacky Pictures and therefore it is relevant and important to assess individual differences in degree of conflict predisposition.

Finally, attention should be called to the fact that the findings have numerous theoretical and clinical implications. The factor intercorrelations bear directly on the syntax of psychoanalytic theory (see Blum, 1949) and the relationships to criterion variables are replete with provocative evidence. It has not been the purpose of this report to pursue such issues, but the interested reader can perform a "do-it-yourself" analysis of many topics from the information presented here. For example, one can piece together all the significant correlations between factors and 1) major fields of specialization in college; 2) physical complaints; 3) sibling constellations; 4) perceived child-rearing practices of family; 5) defense preferences; 6) Strong Voca-

examiner to change from one item to the next rapidly enough to curtail undue deliberation over the alternatives.

tional Interest test; etc. All these separate topics contain intriguing data for theoretical speculation and, especially, follow-up investigations. If enough of the latter materialize, the time and effort put into the development of this research guide will have been well spent.

SUMMARY

This report is the culmination of a lengthy investigation of responses to the Blacky Pictures. All themes, inquiry answers, and cartoon preferences of 210 male undergraduates were recorded, intercorrelated, and factor analyzed separately on each of the eleven pictures. Two or three orthogonal factors emerged per picture, making a total of thirty. Together these comprise a new, improved system for research use of the technique. Scoring instructions are included with the factor descriptions.

Next the factors were related to a host of criterion variables in order to ascertain their construct validity. Available measures provided information on siblings, defense preferences, interests and values, demographic family characteristics, perceived child-rearing practices of parents, past and present physical complaints, field of

specialization in college, grade point average, social perception, interpersonal mechanisms, and others. A summary interpretation of each factor takes into account those findings which achieved statistical significance.

The concluding discussion highlights advantages of the new scoring guide over its predecessors and implications of the results for research on a wide variety of topics.

REFERENCES

- Blum, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.*, 1949, 39, 3-99.
- Blum, G. S. *The Blacky Pictures: a technique for the exploration of personality dynamics*. New York: The Psychological Corporation, 1950.
- Blum, G. S. Revised scoring system for research use of the Blacky Pictures, 1951. Mimeo.
- Blum, G. S. Defense preferences in four countries. *J. Proj. Tech.*, 1956, 20, 33-41.
- Blum, G. S. & Hunt, H. F. The validity of the Blacky Pictures. *Psychol. Bull.*, 1952, 49, 238-250.
- Freedman, M., Leary, T., Ossorio, A., & Coffey, H. The interpersonal dimensions of personality. *J. Pers.*, 1951, 20, 148-162.
- Sirota, L. M. A factor analysis of selected personality domains. Doctoral dissertation, Univ. of Michigan, 1957.

Received August 7, 1961

The Effect of Color on Productivity in Cards VIII, IX, X of the Rorschach¹

JOHN E. EXNER, Jr.
DePauw University

A review of literature pertaining to the stimulus effects of color on Cards VIII, IX, X as a factor influencing a subject's productivity generally reveals that chromatic color has little or no effect in precipitating the number of responses given. The first of such studies was accomplished by Sappenfield and Buker (1949) in which a group administration technique using both chromatic and achromatic slides with normal subjects was reported. They found no differences in productivity. Dubrovner, VonLackum, and Jost (1950) found no differences in productivity or reaction time between the standard and achromatic series given to 30 nurses. The 30 nurses had been previously divided into two groups of 15 each. One group received the standard Rorschach series first and an achromatic facsimile of the Rorschach second, whereas the other group had the order of presentation reversed. Allen, Manne, and Stiff (1951) used the standard and achromatic sets of Rorschach cards with 25 essentially normal college students. They evaluated differences in productivity by using a test-retest method over a six-week interval. Perlman (1951) used the standard Rorschach series and an achromatic facsimile with 80 subjects. She also used the test-retest technique with an interval of one week. Her subjects were divided into four groups, three of which received the achromatic series in the first administration and the standard series in the second administration, whereas the order of presentation was reversed for the fourth group. Her results indicate no differences in productivity on Cards VIII, IX, and X. Meyer (1951) used 60 students in an investigation

of color shock in the Rorschach. He divided these students into two groups of 30 each, matching his subjects on the basis of age and sex. While his report is mainly concerned with the various indices of color shock, he incidentally reports no difference in productivity for Cards VIII, IX, and X.

While each of the studies cited generally concludes that color, as a stimulus characteristic is not significantly influential in productivity, it would appear that additional investigation into this problem area is necessary since each of the studies has distinctive problems in experimental design which may have loaded the procedure in favor of the findings derived. Sappenfield and Buker; Allen, Manne, and Stiff; and Perlman all used a test-retest technique. It seems obvious that the naïveté of the subject would be lost in the retest session and thus a set would have been developed with regard to the general test procedure and more specifically with regard to the form, shading, and general structure of the blots so that productivity might well not be expected to increase. Dubrovner, VonLackum, and Jost also used the test-retest technique and may have additionally compounded this experimental error by selecting subjects all of whom lived together, thus increasing the possibilities that the subjects were not naïve at the time of the testing. Meyer's experiment relates most closely to the present design. One oversight which may have occurred in his study is in the area of matching of his subjects. While they were matched on the basis of age and sex, they were not matched on general intellectual ability. While all of his subjects were college students, the differences in intellectual potential and especially in being able to articulate responses could feasibly

¹ This research was supported in part by a grant from the Graduate Research Council, DePauw University.

have been significantly different in his small groups so as to precipitate the types of results derived.

The present investigation was designed to deal specifically with the stimulus characteristics of chromatic color on Cards VIII, IX, and X as an element influencing response productivity. It was hypothesized that the chromatic color as an independent variable would elicit more responses on each of Cards VIII, IX, X of the Rorschach. This hypothesis in part was arrived at through a previously reported study by Exner (1959) dealing with Card I of the Rorschach. In that investigation it was demonstrated that when Card I was altered to have chromatic rather than achromatic color the number of responses given increased significantly over the number of responses given by a matched group to the standard Card I of the Rorschach series.

PROCEDURE

Forty college students were selected to be used as subjects. They were then divided into two groups of 20 each. Subjects in one group were matched with the subjects in the second group on the basis of age, sex, marital status, academic class standing, verbal score on the Scholastic Aptitude Test, and cumulative grade-point average over at least three semesters. The matching was as exact as possible so that all factors were held absolutely constant with the exception of grade-point average which varied only slightly, at the most 1/10 of one point on a three-point scale.

The materials used in the experiment were the standard Rorschach Cards VIII, IX, X and achromatic facsimiles produced photographically using a 1:1 negative. The prints were made on a glossy bromide paper similar to that used for the Rorschach and each negative was printed until the lightest area of print matched the shading on the original cards. In this way form, background, size and shad-

ing were held as constant as possible.²

The test procedure was such that each subject was tested individually using standard Rorschach instructions except that the test began with Card VIII and proceeded through Cards IX and X. The inquiry following the free-association period was limited only to a determination of location of response rather than determinants to expedite the test process and also for the fact that such different stimulus characteristics were involved that the determinants could not easily be compared. The standard chromatic VIII, IX, X cards were administered to all subjects in Group I, whereas the achromatic facsimiles were administered to all subjects in Group II. No retesting was involved.

RESULTS

The total number of responses by each group to each of the three cards is shown in Table I. These data were analyzed using a matched *t* test.

TABLE I. Total Number of Responses Per Card

Card	Group I Chromatic	Group II Achromatic	<i>t</i> Value
VIII	49	33	2.96**
IX	45	29	2.67**
X	72	50	2.08*

*represents statistical significance beyond .05.

**represents statistical significance beyond .01.

Examination of Table I indicates that there were consistently more responses given to the chromatic cards VIII, IX, X than to the achromatic facsimiles. The differences between the number of responses for each of the three cards and for the total number of responses is statistically significant for Cards VIII and IX beyond the .01 level and for Card X beyond the .05 level.

The data presented in Table II represent the mean reaction times for

² The achromatic cards used in the experiment were reproduced by the Graphic Arts Research Division of Rochester Institute of Technology under the supervision of Mr. Lester Stroebel.

first responses to each card. These data were analyzed using a matched *t* test.

a tendency toward *W* responses rather than *D* responses.

TABLE II. Mean Reaction Times for First Responses to Each Card

Card	Group I Chromatic	Group II Achromatic	<i>t</i> Value
VIII	25.25	17.10	1.23
IX	48.80	30.80	2.38*
X	27.65	26.15	0.03

* represents statistical significance at .02

Examination of Table II indicates that there was a statistically significant difference between the mean reaction times to Card IX in that more time was required to respond to the chromatic card than to the achromatic facsimile. No significant differences in mean reaction times occurred to either of the other cards.

The data presented in Table III represent the number of *W* and *D* responses given to each card by each of the groups. These data were analyzed using a proportional chi-square.

Examination of Table III reveals that a significantly greater proportion of *W* responses occurred to the achromatic Card VIII than to the chromatic Card VIII and that a significantly greater proportion of *D* responses occurred to the chromatic Card X than to the achromatic Card X.

Statistical analysis was also made concerning the two popular responses which often occurred to these three cards; that is, the response of an animal on Card VIII and a spider or crab to Card X. This analysis reveals no significant differences; however, there is a considerable decrease in the number of animal responses to Card VIII, probably having been created by

DISCUSSION

The results of this experiment seem to give considerable weight to the assumption that color does affect productivity. It is difficult to relate this phenomena to the various theories concerning color; however, it would appear that Klopfer's original hypothesis concerning a 8-9-10 per cent (1942) is not as erroneous as previous studies have led to believe. Obviously a more thorough investigation into the usefulness of the 8-9-10 per cent might be accomplished wherein a complete Rorschach series and an achromatic facsimile of a complete series would be administered to subjects in two different but highly matched groups. The results from such an experiment could then give some realistic indication as to the correctness of Klopfer's assumption that 30 per cent of the total number of responses in a record should be given to Cards VIII, IX, and X.

Possibly a more enlightening factor evolving from this investigation seems to pertain to the experimental design. Rorschach investigators traditionally use small samples, since individual administration with large groups becomes somewhat unwieldy. Altogether too often these small samples are apparently randomly grouped rather than being grouped on a highly selective basis. Many studies in perception have demonstrated that intellectual capacity, especially as related to abstract thinking, and ability to be articulate play highly important roles in the responsiveness of subjects. Yet a variety of so-called "validation" studies have been accomplished where

TABLE III—Total Number of *W* and *D* Responses Per Card

Card	Group I Chromatic	Group II Achromatic	Proportional Chi ² Value	Group I Chromatic	Group II Achromatic	Proportional Chi ² Value
	<i>W</i>	<i>W</i>		<i>D</i>	<i>D</i>	
VIII	28	24	8.76*	22	10	0.10
IX	20	14	0.46	24	12	0.88
X	24	17	1.12	46	31	10.40*

* represents statistical significance beyond .01.

these factors have been minimized or completely neglected. In this study no effort was made to match for articulate ability; however, the exactness of matching of intellectual potential and demonstrated academic achievement seems to have played an important role, especially when the results are compared with those obtained by Meyer in his study of a similar design.

Since clinical psychology is intensely concerned with demonstrating the reliability and validity of projective techniques, equal concern must be given to the experimental designs by which these studies are undertaken. There is constant discussion pertaining to individual differences, and it has not been uncommon for clinicians to view negative results pertaining to the reliability or validity of projective techniques with some caustic comment such as "Yes, that may hold true for that group but there are always individual differences which must be considered." While it is true that individual differences must be considered, they can be overcome to some extent by exact matching of subjects. If this is accomplished in future studies, especially those which pertain to the perceptual aspects of projective materials, the results may be viewed with more confidence and the end product will be a greater fund of knowledge concerning the devices so commonly used.

SUMMARY

Forty college students were divided into two groups of 20 each. Subjects were selected on an exact matching

basis using the criteria of age, sex, marital status, intellectual capacity, and academic grade point. Subjects in one group were individually administered Cards VIII, IX, and X of the standard Rorschach series, while subjects in the second group received achromatic facsimiles of these cards. Results indicate a significantly greater number of responses to each of the chromatic cards than to the achromatic facsimiles, thus suggesting that color does affect productivity on Cards VIII, IX, and X of the Rorschach. Mention was made of the necessity for exact matching of subjects in perception studies pertaining to projective techniques.

REFERENCES

- Allen, R. M., Manne, S. H. & Stiff, M. The role of color in Rorschach's test. *J. proj. Tech.*, 1951, 15, 235-242.
- Dubrovner, R. J., VonLackum, W. J. & Jost, H. A study of the effect of color on productivity and reaction time in the Rorschach test. *J. clin. Psychol.*, 1950, 6, 331-336.
- Exner, J. E. The influence of chromatic and achromatic color in the Rorschach. *J. proj. Tech.*, 1959, 23, 418-425.
- Klopfer, B., & Kelley, D. M. *The Rorschach technique*, Yonkers, New York: World Book, 1942.
- Meyer, B. T. An investigation of color shock in the Rorschach test. *J. clin. Psychol.*, 1951, 7, 367-370.
- Perlman, Janet. Color and the validity of the Rorschach 8-9-10%. *J. consult. Psychol.*, 1951, 15, 122-126.
- Sappenfield, B. R., & Buker, S. D. Validity of the Rorschach 8-9-10%. *J. consult. Psychol.*, 1949, 13, 268-271.

Received July 19, 1961

Levels of Awareness in Projective Tests¹

GORDON FILMER-BENNETT
Winnebago State Hospital

WALTER G. KLOPPER
University of Portland

In making predictions from various projective tests, confusion frequently arises concerning the level of conscious awareness being tapped. This is especially true as regards comparisons of sentence-completion and Thematic Apperception Test protocols. These latter issue from test stimuli often considered to be of proportionately equivalent structure in contrast to that of certain other projective tests such as the Rorschach or Draw-A-Person tests. Hanfmann (1947) has contended that the SCT more closely approximates the manifest level than does the TAT. A recent study by Stone and Dellis (1960) has offered token support to this claim by uncovering a less-than-significant difference between the SCT and TAT in rated amount of pathology elicited. At the same time, since clinical interpretation often enough treats the data from these two tests as though they measure the same behavioral level, the matter merits further investigation.

The object of the present study was to determine predictability of self-ratings from SCT and TAT protocols. Considered within the framework of Leary's (1957) multi-level theory, the problem may be stated in terms of the correlation of Level II (Conscious Communication) and Level III (Private Perception) behavior.

METHOD

Nine TAT cards (I, II, IIIGF, VIIIGF, IV, VIIIBM, IXGF, XVIIIGF, and X) and a form of the Forer Sentence-Completion Test were individually administered by the same examiner to 20 female junior college

students, using standard instructions. Half of the subjects took the TAT immediately prior to the SCT, while for the remaining subjects the order was reversed. The protocols were then randomly arranged and assigned code numbers. The series concluded with a self-rating scale which required that the subjects rate themselves on each of four traits along a seven-point continuum. These traits were pretested and selected from among seven original traits on the basis of their approximation to a normal distribution. The resulting continua included the following:

- T1: Communicative-uncommunicative
- T2: Intellectual-nonintellectual
- T3: Conventional-unconventional
- T4: Sensitive-insensitive

Three experienced clinical psychologists were then requested to Q-sort the subjects on the basis of their TAT performance according to how each would be likely to rate herself on each of the aforementioned seven-point continua. An additional three psychologists were asked to do a similar Q-sort based upon the SCT protocols. In each instance the Q-sort entailed forced ranking to conform to a normal distribution.

RESULTS AND DISCUSSION

Tables I and II list the rank-difference coefficients describing the relationship between self-ratings and judges' ratings on each of the traits. Inspection suggests that the only significant predictions are those based upon TAT data associated with communicability, all other predictions being of uniformly lower or negligible value. It is noteworthy that inter-judge reliability was also consistently higher

¹Appreciation is extended to Dean W. E. McDonald and the student participants of Norfolk Junior College, and especially to the six judges who gave so liberally of their time.

TABLE I. Relationship of Judges A, B, and C's SCT Ratings with Self-ratings on Each of Four Traits, Expressed as Rank-Difference Coefficients

Self-rating	T1			T2			T3			T4		
	A	B	C	A	B	C	A	B	C	A	B	C
T1	.27	.09	-.15									
T2				.13	.27	.27						
T3							.07	-.29	.07			
T4										-.04	.07	.11

TABLE II. Relationship of Judges A, B, and C's TAT Ratings with Self-ratings on Each of Four Traits, Expressed as Rank-difference Coefficients

Self-rating	T1			T2			T3			T4		
	A	B	C	A	B	C	A	B	C	A	B	C
T1	.62	.86	.76									
T2				-.02	-.30	-.34						
T3							.29	.47	-.05			
T4										.01	.06	.02

($\rho = .56, .67, .71$) in the case of these particular TAT predictions than in all other instances. These data together suggest that sheer productivity on the TAT protocols may have provided a more "objective" basis for predicting communicability, one on which the judges could more readily agree, whereas determination of the other traits may have utilized more intangible, intuitive criteria.

In general, the tabulated findings provide some basis for concluding that the SCT and TAT samplings may have been, for the most part, psychologically equidistant from Level II behavior. At the same time, a spot check suggests that negligible correlations prevail between the SCT ratings on the one hand and the TAT ratings on the other, implying that information deriving from these two tests is more apt to be mutually supplementary than overlapping in nature insofar as present purposes are concerned. An alternative possibility is that the SCT and TAT may in some respects be inadequate measures of the Level II traits in question. Here cognizance may be taken of McGreevey's (1959)

suggestion that failure in predicting Level II from Level III behavior may be partly a function of the ego-threatening nature of the trait being measured, even though in the present instance an effort was made to select those traits least likely to provoke defensiveness.

CONCLUSIONS

It has been widely assumed that the SCT taps behavior which is closer to awareness than does the TAT. The results of this study cast considerable doubt upon this assumption.

REFERENCES

- Hanfmann, Eugenia. Projective techniques in the assessment program of the Office of Strategic Services. In *Exploring individual differences*. Washington, D.C.: Amer. Council on Educ., 1947, pp. 19-29.
- Leary, T. *Interpersonal diagnosis of personality*. New York: Ronald Press Co., 1957.
- McGreevey, J. C. The effect of inter-level disparity on predictive efficiency. Portland, Ore.: University of Portland, 1959. (Unpublished doctoral dissertation)
- Stone, H. K. & Dellis, N. P. An exploratory investigation into the levels hypothesis. *J. proj. Tech.*, 1960, 24, 333-340.

Received May 18, 1961

Factors Influencing Rotation In the Bender-Gestalt Performance of Children¹

JERRY B. FULLER² AND GILLES CHAGNON
University of Ottawa

INTRODUCTION

During the performance of the Bender-Gestalt test many children reproduce the designs correctly but in a rotated position (Baroff, 1957; Bender, 1952; Bensberg, 1952; Byrd, 1956; Clawson, 1959; Feldman, 1953; Goldberg, 1957). Studies on the Bender-Gestalt test fail to explain adequately why this distortion in reproduction occurs. Hanvik (1953) found a relationship between rotation and brain-injury in children. In children with learning problems rotation was ascribed to regression (Fabian, 1945; Koppitz, 1958). This study attempts to provide a tentative rationale for rotation on the Bender-Gestalt cards based on the influence of stimulus variation, cue utilization or a combination of both.

THEORETICAL CONSIDERATIONS

Boring (1952, p. 144) in a recent article stated: "The effective stimulus is not an object but a property of the stimulus object, some crucial property that cannot be altered without changing the response, some property that remains invariant, of other characteristics". In this study, it is postulated that the Bender cards and their designs, have a number of such properties which may not necessarily be of equal importance in producing rotation. Systematic variation of such properties as figure, ground and configurations should throw light on their influence upon rotation. These variations will be referred to as "sets".

¹ This article is based on a dissertation submitted by the senior author to the Department of Psychology of the University of Ottawa in partial fulfillment for the degree of Doctor of Philosophy.

² Now at Willmar State Hospital, Minnesota.

Figure

The word figure in this study applies to the design which the subject has to reproduce. It can be either horizontally or vertically orientated to the background. The five original Bender figures used in this study are all in a horizontal relationship to the background. That is, both figure and background are orientated in the same direction. Studies using the Kohs Blocks Goldstein and Scheerer (1941) and the Block Design Test (Shapiro, 1952; Yates, 1956); have shown that more rotation is produced when figure and ground are incongruent; that is, when figure is at an angle to the axis of the background than when the figure is parallel to it. The results of these studies were further supported by Bakay and Schiller (1948). If orientation of a figure thus affects the subject's performance, it can then be hypothesized that rotation will occur more frequently on the Bender cards when the figure is in a vertical position than in a horizontal one.

Ground

The ground defined as the white card on which the design is placed, has been shown to affect the amount of rotation. Studies using the Cube Test Goldstein and Scheerer (1941, p. 39-43) Kohs Blocks (Shapiro, 1952; Yates, 1956) and Block Design Test Williams et al (1956) have presented evidence to show that a diamond-orientated ground produced more rotation than a square-orientated ground. The above studies give some indication that the properties of a diamond figure are not as substantial as those of a square figure.

Can it be assumed that the properties of a rectangular shaped background, such as a Bender card, would produce the same, more, or less rotation than when presented in a diamond shaped one? Helson and Fehrer (1932) in a quantitative study concerning 3 kinds of criteria for preference of different types of form showed that different criteria favored different forms. The figures used were the triangle, angle, square, circle, rectangle and diamond. They concluded: "If we were to choose the best form on the basis of all our criteria taken together, the rectangle would get first place". This study was further verified by Casperson (1950), who used five kinds of criteria and found that the rectangle and square surpassed the diamond in every case. In the above studies, there is evidence that a diamond-shaped ground tends to produce more rotation than a square-shaped ground; secondly, that in relation to a diamond-shape object a rectangular shape is, in terms of form, as 'good' and as 'simple' in structure as a square.

Therefore, it might be hypothesized that rotation will occur more often on the Bender cards when the background is in the shape of a diamond than when it is in the shape of a rectangle.

Configurations

The configurations used in this experiment were designs A, 1, 2, 3 and 8 of the original Bender designs (1938). These configurations were selected because they have been found to produce more rotation than the other Bender figures (Goldberg, 1957; Koppitz, 1958; McPherson, 1955). In consideration of each individual configuration it has been pointed out that they are not equivalent items and that different configurations would elicit different responses in regard to distortion (Baroff, 1957; Byrd, 1956; Goldberg, 1957; Koppitz, 1958; McPherson, 1955). Therefore, it can be

hypothesized that some of the configurations produce more rotation than others.

Cue Utilization

The use of cues or, as designated in this study, the range of cue utilization, was described in general by Easterbrook (1959, p. 183) as, "the total number of environmental cues in any situation than an organism observes, maintains an orientation towards, responds to, or associates with a response". Cue utilization is equivalent to the amount of information in use by a person at any one time. The information used is related to the number of cues employed. Therefore, the more difficult the task, the greater amount of information or number of cues needed (Bahnick, 1954; Broadhurst, 1959). It has been shown that the range of cue utilization may be thought of as "total information used" (Born and Rubenstein, 1952; Hick, 1952; Miller, 1956). It is then felt that an individual is capable of an adequate performance when he is able to use all the cues or information available at any one time for an appropriate response.

It has further been demonstrated by Easterbrook (1959), Duffy (1948), Calloway and co-workers (1958) that: a) emotions tend to reduce this range of cue utilization, b) the attentive field tends to become narrower in emotionally disturbed subjects and c) the perceptual field tends to be reduced.

Yates (1954) showed that subjects are able to counteract rotation producing stimuli, provided the usual visual indications of 'verticality' and 'horizontal' are available. However, when these indications were artificially removed many normal subjects began to rotate to a considerable extent, since they were then more dependent on the rotation inducing cues.

In a study on space orientation,

Witkin and Asch (1948) found that the directional qualities perceived by subjects are a function of the organization of stimulation available to the subjects at that time. It was further indicated that many of the subjects' perceptions of a field were functions of the visual, directional cues available. In another study by Asch and Witkin (1948) it was demonstrated that errors in the perception of the vertical field varied, according to the amount of information derived from the surroundings.

It follows from the foregoing studies that emotions tend to reduce the range of utilization and it is further felt that rotation may be influenced by a lack of cue utilization due to emotional excitation, arousal or disturbance. The degree of emotional excitation, arousal or disturbance, hence range of cue utilization is assumed in this experiment to be isotonic to the difference in degree of disturbance of the three groups of individuals used as subjects. Thus, for convenience sake, range of cue utilization will be designated under the term "groups".

From the aforementioned rationale and considerations the following general hypothesis can be set forth: Rotation will be influenced differently by the six "sets" depending upon the figure-ground orientation; the five "configurations" will influence rotation in different ways because they are not equivalent items and the three "groups" will influence rotation differently according to their availability of range of cue utilization.

For the sake of experimental verification the general hypothesis was broken down into the following sub-hypothesis:

1. There is no significant difference between the amount of rotation produced by any two of the six sets of cards.

2. There is no significant difference

between the amount of rotation produced by any two of the five configurations.

3. There is no significant difference between the amount of rotation produced by any two of the three groups.

4. There is no significant interaction between the six sets and the five configurations in the amount of rotation produced.

5. There is no significant interaction between the six sets and the three groups in the amount of rotation produced.

6. There is no significant interaction between the three groups and the five configurations in the amount of rotation produced.

7. There is no significant interaction among the six sets, five configurations and the three groups in the amount of rotation produced.

SAMPLE

Three sample groups were used in this study. These were normal, disturbed and schizophrenic subjects between eight and fifteen years of age. The upper age limit of fifteen was chosen because patients above this age are considered adults rather than children for purposes of diagnosis and treatment. The lower age level of eight was selected because it has been shown (Bender, 1938; Fabian, 1945; Koppitz, 1958) that rotation is expected in normal children up to the age of seven and that children under eight cannot reproduce the Bender figures adequately.

The normal subjects were ninety children of average or above average intelligence who had never been referred for psychiatric help nor considered behavior problems by school authorities, teachers or the school psychologist.

The second group consisted of ninety children who were not schizophrenic, mentally defective or known to

have brain lesions at the time of testing selected at random from various child guidance clinics. Many syndromes were present, but they can all be subsumed under the term: maladjusted behavior associated with emotional disturbances. All children at the time of testing had been referred to a clinic for psychological assistance. The records were checked to make sure there was no gross visual field defect present.

Ninety schizophrenic children made up the last group. They were diagnosed as such by the psychological staffs in the hospitals in which they were found. All children were excluded who were known to have detectable brain damage, mental deficiency or an obvious physical handicap. A double check of the diagnosis was made in a majority of the cases. An out-patient's clinic had originally diagnosed as schizophrenic most of these children and had referred them to other facilities. With the list of children available from the out-patient clinic, it was possible to re-check the diagnosis through the institution in which they were presently receiving care.

With an age range of eight to fifteen for the three groups, the mean age of the normal children was 11.54, the emotionally disturbed children, 11.24 and the schizophrenic children, 11.70.

INSTRUMENT

The test consists, essentially, of thirty white cards, fifteen of the cards measuring four by six inches and the remainder measuring five by five inches in size. On each of these cards one of the five Bender configurations was printed. The configurations themselves were numbered one to five^a. The thirty cards of the test consisted, therefore, of six sets of five cards, each set

consisting one each of the five configurations. These six sets were lettered a, b, c, d, e and f and were arranged in the following manner:

1. Set a consisted of the original Bender configurations and cards, with rectangular ground and horizontal configurations.
2. In set c all configurations were rotated ninety degrees counterclockwise to produce the vertical effect. The ground shape remained rectangular.
3. The ground shape in set e was in the form of a diamond with the configurations horizontal.
4. Set f was a combination of sets c and e. The configuration was in a vertical position and the ground was in the form of a diamond.
5. In set d the rectangular card was rotated ninety degrees counterclockwise to give a vertical ground effect, with the configurations remaining in a horizontal position.
6. Set b was a combination of sets c and d. The configurations were rotated ninety degrees counterclockwise so that they would be in a vertical position and the rectangular card was rotated ninety degrees so that it would be in a vertical position.

The relationship obtained may be seen more clearly by referring to figure 1, which presents a concrete illustration of each type of combination.

Rotating of the cards and figures ninety degrees counterclockwise was an arbitrary decision on the part of the writers. They could have been just as easily rotated ninety degrees clockwise. This technique was used to assure uniformity throughout the sets.

ADMINISTRATION

The test was individually administered within the three groups. Each individual had to draw three configurations with only the last configuration being scored for degree of rotation. The first two cards, a square and a triangle were in addition to the thirty

^a Figure 1 of the present study corresponds to figure A of Bender's standard series. Figure 2 corresponds to figure 1; figure 3, to figure 3; figure 4 to figure 2; figure 5 to Bender's figure 8.

TABLE 1 Summary of Analysis of Variance of the 6 x 5 x 3 Factorial Design.

Source of Variation	SS	df	Estimate of Variance	F
G	23484.33	2	11742.16	29.80*
(Groups)				
S	9131.52	5	1826.30	4.63*
(Sets)				
C	6127.28	4	1531.82	3.88**
(Configurations)				
G x S	20659.77	10	2065.97	5.24*
G x C	6304.38	8	788.04	1.96*
S x C	24318.24	20	1215.91	3.08*
G x S x C	14260.63	40	356.51	
Intra	70929.67	180	394.05	

* significant beyond .001

** significant beyond .01

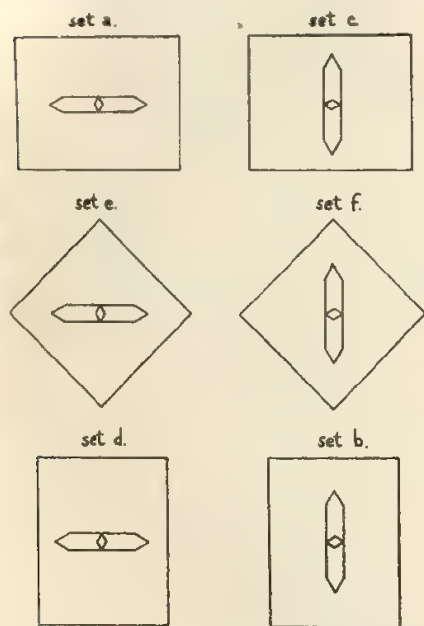


FIG. 1 An Example of the Different Combinations of Figure and Ground for the Six Sets to be Used in This Study.

cards used in this experiment. They served as a buffer or practice cards and were given to all subjects. Since each subject would only be requested to draw one configuration it was felt that they should have the opportunity to warm up or get accustomed to this type of testing situation. Individuals within a group were assigned at random to each of the thirty cards.

SCORING

Rotation was considered as any reproduction of a drawing which altered the actual axis of the drawing. This could only be produced in one way; the subject rotated the configuration while copying it from a properly oriented piece of paper. The score recorded for an individual was the total number of degrees by which the drawing deviated from its original axis.⁴

STATISTICAL ANALYSIS

The general statistical approach used in this study is that of a triple classification factorial design, employing analysis of variance. The three variables used in the present experiment were sets, configurations and groups.

The model used was the ABC type in which only systematic variation is taken into consideration.

The residual mean square within groups was chosen as the error term for this type of model since it is generally considered the appropriate term to use.

THE RESULTS OF THE EXPERIMENT

The principal results obtained with respect to the hypotheses put forward in this experiment may be summarized as follows:

⁴ Full details of the method of administering, scoring and measurement of rotation are contained in Fuller (1960).

Sets

The F-test for rotation scores for the main effect, sets, as seen in Table I is significant beyond the .001 per cent level of confidence, and the null hypothesis can be rejected.

T-tests were used to determine the significance between the means of the sets. Results of the t-test on differences in the means of rotation scores for sets are given in Table II.

It will be noted from Table II that set a is significantly different from sets c, d, e and f, and that set b is significantly different from sets d, e and f. Sets a and b produced very little rotation when compared to sets c, d, e and f. In fact, it would make little difference which of the last four sets was

used in this study since each seems to contribute about the same amount of rotation. This is why the differences between any of the other combinations were not significant.

Configurations

It can be seen in Table I that the F test for rotation scores for the five configurations yielded significant values at the .01 per cent level of confidence, and the null hypothesis can be rejected. The t-test to determine the significance of the difference between the means of rotation scores for configurations was applied. From Table III it may be seen that configuration three was significantly different from configurations two, four and five and is consistently producing a high

TABLE II Evaluation of the Difference of Means on the Rotation Scores of the Six Sets by Means of the t-Test.

Sets Used	Mean Values Compared	Difference	Significant Difference
a-b	3.20- 4.30	1.10	No
a-c	3.20-14.20	11.00	Yes*
a-d	3.20-17.30	14.10	Yes
a-e	3.20-23.40	20.20	Yes
a-f	3.20-20.60	17.40	Yes
b-c	4.30-14.20	9.90	No
b-d	4.30-17.30	13.00	Yes
b-e	4.30-23.40	19.10	Yes
b-f	4.30-20.60	16.30	Yes
c-d	14.20-17.30	3.10	No
c-e	14.20-23.40	9.20	No
c-f	17.30-20.60	6.40	No
d-e	17.30-23.40	6.10	No
d-f	17.30-20.60	2.30	No
e-f	23.40-20.60	2.80	No

* t ($p = .01$), significant difference ≥ 10.37

TABLE III Evaluation of the Difference of Means of the Rotation Scores of the Five Configurations by Means of the t-Test.

Configuration Used	Mean Values Compared	Difference	Significant Difference
1-2	14.40-11.30	3.10	No
1-3	14.40-22.50	8.15	No
1-4	14.40- 8.50	5.90	No
1-5	14.40-12.40	2.00	No
2-3	11.30-22.55	11.25	Yes*
2-4	11.30- 8.50	2.80	No
2-5	11.30-12.40	1.10	No
3-4	22.55- 8.50	14.05	Yes
3-5	22.55-12.40	10.15	Yes
4-5	8.50-12.40	3.90	No

* t ($p = .01$), significant difference ≥ 9.92

TABLE IV—Evaluation of the Differences of Means on the Total Rotation Scores of the Three Groups by Means of the t-Test

Groups Used	Mean Values Compared	Difference	Significant Difference
Normal-Emotional	190-15.00	13.10	Yes*
Normal-Schizophrenic	190-24.62	29.72	Yes
Emotional-Schizophrenic	15.00-24.62	9.62	Yes

*t ($p = .01$), significant difference ≥ 7.88

degree of rotation in comparison to other configurations. No significance was found between the other configurations.

Groups

Table I shows that the F-test for the main effect, "groups", is significant beyond the .001 per cent of confidence. Therefore the null hypothesis can be rejected. T-tests were applied to find where the specific differences occurred. From Table IV it can be seen that all three groups differ from each other significantly in total mean rotation scores.

Sets by Groups

The test for significance of double interaction, sets by groups, resulted in a significant interaction. The F-test was significant at the .001 per cent level of confidence. Therefore, the null hypothesis was rejected. As previously, t-tests were applied to find where the differences occurred.⁵ From the data it can be seen that the normals produced the least rotation and the schizophrenics the most. Also, the emotionally disturbed group's scores on all six sets differ significantly from those of the schizophrenics on set f. Both the emotionally disturbed and schizophrenic groups performances on sets c, d, e and f were significantly different in mean rotation score from that of the normal group on sets a, b, c, d and e.

Sets by Configurations

The F-test demonstrates a significant interaction at the .001 per cent level of confidence. Therefore the null

hypothesis was rejected. T-tests were used to determine where the significance between the means of the sets by configuration was to be found. The analysis shows that set d, configuration 3, was significantly different from all the other sets for all the configurations. The next best combination was set 3, configuration 3, followed by set f, configuration 1 and 5.

Groups by Configurations

For the double interaction, groups by configuration, the F-test was not significant and the null hypothesis cannot be rejected.

Sets by Configurations by Groups

For the triple interaction, sets by configurations by groups, the F-test was not found to be significant. The null hypothesis was not rejected.

DISCUSSION OF THE DATA

Sets

In the last section it was shown that sets c, d, e and f produced the most rotation. This supports the rationale that certain stimulus variations of the Bender test in terms of variation of figure and ground will produce more rotation than others. It may be concluded, furthermore, that a vertically orientated figure produces more rotation than a horizontally orientated figure when the same ground shape is used as long as there is incongruency between figure and ground. When the ground shape was vertical and incongruent to the figure there was even more rotation produced. A diamond orientated ground produced more rotation than a rectangular orientated ground regardless of the position of

⁵ Copies of the tables of double interactions may be obtained upon request.

the figure. A diamond oriented ground when incongruent to the figure shape produced more rotation than when it was not.

Configurations

It was found that configuration 3 contributed significantly more to rotation than the other four configurations. Configuration 1 followed closely behind configuration 5. This observation lends support to the rationale that the configurations are not equivalent items and therefore will contribute differently to the amount of rotation produced. Looking at configuration 1, it might contribute more to production of rotation because an individual has to be able to integrate both parts of the design into a whole gestalt. If a subject cannot handle the configuration as a whole unit, he will deal with the configurations as if they were isolated details which might make proper orientation of the configuration more difficult. On the other hand, since a diamond has poorer form than a circle, an individual might be distracted by either the stronger form or the weaker one, thus making it harder to orientate the configuration. This seems to stress the importance of the relative strength of competing perceptual organization. It's the complexity and competing forces of form which increases the amount of rotation.

For configuration 3 which is based on the principle of proximity of parts, it seems that the very structure of the configuration is the basis for it producing more rotation. It has been shown by Bender (1938) that this is the most difficult configuration to reproduce and is not fully mastered until age eleven. In other words, the subject might have trouble encompassing the total configuration and instead deals with the separate units which are segregated and yet at the same time belong to a larger unit. The effects of the meaning of a stimulus in terms of the attitudes and values of the observ-

er appeared to be important. Many of the subjects interpreted this configuration as being a christmas tree and reproduced it in light of such interpretation which caused it to be rotated. It appears that proper orientation and cue availability were subsumed in favor of personal meaning.

Groups

The significant differences found between the three groups in mean rotation scores supports the rationale that rotation is influenced by a lack of cue utilization due to emotional excitation, arousal or disturbance. That is, the amount of rotation which would take place is at least to some extent a function of the conflict between those factors inducing rotation (configurations and sets) and those factors working against rotation (utilization of visual directional cues in the groups). It follows, therefore, that the more disturbed an individual is, the fewer visual cues he will have available to orientate himself to those sets producing rotation.

The emotionally disturbed and schizophrenic groups, being more disturbed would be less aware of the surrounding stimuli that are important for visual cues. When compared to the normal group they should rotate more because there will be available to them fewer rotation-counteracting cues and hence they will be more influenced by the rotation producing sets.

Sets by Groups

Earlier it was shown that the interaction of sets by groups was significant. This indicates that not only were the groups, sets and configurations contributing to production of rotation individually but a combination of two or more of these factors was influencing rotation, the first combination being groups by sets. Therefore, cue utilization and figure-ground orientation both were operating in production of rotation.

The assumption that the normal

group is free of emotional excitation, arousal or disturbance and therefore is able to use all the cues necessary at any one time for an appropriate response seems tenable. Indeed there seemed to be no conflict between the rotation producing sets, regardless of their figure-ground orientation and factors working against rotation such as availability of cues for this group. It would seem that the normal subjects were able to counteract the rotation-influencing stimuli (diamond shaped ground and vertical figure and rectangular vertical ground orientation) because the necessary visual cues were available to them and used by them. This is brought out even more clearly by the fact that the two disturbed groups differed significantly in their mean rotation scores on sets c, d, e and f which are found to produce more rotation, when they are compared to the normal groups scores on all the sets.

This means that the two disturbed groups were affected both by the figure-ground orientation and the amount of cue utilization used. They were not able to handle or cope with the conflict between the rotation producing sets and factors working against rotation such as utilization of cues. Taking into consideration their disturbance and the fact that a more difficult set to orientate is facing them, it might be possible for them to become more aroused and excited. Therefore making it more difficult for them to draw on enough of the cues available to counteract the rotation producing sets.

In regard to the emotionally disturbed groups, sets c, d, e and f resulted in the greatest rotation with set e (the diamond orientated ground and horizontal figure) producing the most.

In the case of the schizophrenic groups sets c, d, e and f, again, produced rotation with set f producing significantly more rotation than all six sets for the emotionally disturbed

group or the normal group. For the schizophrenic group, then, the diamond orientated ground and vertical figure was the combination which produced the most rotation.

In summary, then, the most striking phenomenon is the absence in the normal group of any influence by the figure-ground orientation or utilization of cues. By contrast the most powerful influences in the emotionally disturbed and schizophrenic groups which produced the rotation effect was firstly the orientation of the ground shape. When the ground shape was a diamond, a great deal of rotation was produced; when the ground shape was a rectangle, very little was produced. The exception is when the rectangle was in a vertical position. Secondly, when the figure shape is incongruent to the ground shape more rotation is produced than when it is congruent. The exception is when there was a diamond ground and vertical figure.

Sets by Configuration

In this combination a significant interaction was found. This means that both figure-ground orientation and configurations together were operating to produce rotation.

For configuration 1 most rotation was produced when used with sets c and f. The factor common to both sets was the placement of configuration 1 in the vertical position.

The greatest amount of rotation was produced when configuration 3 was combined with sets c, d and e. The one factor that stands out in these sets is that the figure is always incongruent to the ground. It should be pointed out that set d produced significantly the most rotation in relation to all the other sets when using configuration 3. In set d the figure-ground orientation was a rectangular shaped ground in the vertical position and the figure in the horizontal position. Therefore, to avoid rotation a subject must have been able to handle the structure of

the most difficult configuration and at the same time the verticalization effect. The other important point is that all the configurations produce more rotation when combined with either set e or f.

The next consideration should be in the direction of a cross validation study which is in progress now. This would mean taking sets d, e and f combined with configurations 1 and 3 and administering them to the same kind of population used in this study to see if consistent results would be found. If significant results were found, the next step would be to establish a cutoff score for rotation for the three groups. Other pathological groups besides the two used in this study could also be administered the new sets to see what kind of results would be obtained, and ultimately produce a more refined diagnostic tool for screening pathological groups.

SUMMARY

This study was concerned primarily with the rotation effect on the Bender-Gestalt performance of children. An attempt was made to put forth some of the underlying rationale to account for this rotation effect. The rationale proposed states that rotation might be influenced by stimulus variation, cue utilization or a combination of the two. Stimulus variation was achieved by the manipulation of a) figure and ground orientational relationships (vertical or horizontal) and shape of ground (diamond or rectangular). The six combinations were designated by the term sets. b) The figures themselves, five of the original Bender figures called configurations. Range or cue utilization considered to be influenced by emotional excitation, arousal or disturbance and varied by using ninety normal, ninety neurotic and ninety schizophrenic children between the ages of eight and fifteen.

The instrument used consisted of six sets of five cards each. The first set contained the original Bender figures

a, 1, 2, 3, 8 and cards and the remaining five sets were varied according to the vertical-horizontal orientations of figure or ground and diamond-rectangular ground shapes.

The main hypothesis to be tested was that there were no significant effects on rotation scores attributable to the groups, the sets, the configurations or a combination of two or more of the above factors.

Interpretation of the results showed that there was a significant difference among the groups which suggests that the more emotionally disturbed, excited or aroused an individual is the less he is able to draw on the necessary cues to avoid rotation. There was significance found among the sets with four of the six combinations producing more rotation. It was found that a vertically orientated figure produced more rotation than a horizontal one and the diamond shape produced more rotation than the rectangular shape. For configurations, it was found that configuration 3 produced the most rotation followed closely by configuration 1.

The differences among groups by sets combinations were statistically significant which shows that both the availability of cues and figure-ground orientation were operating in producing more rotation.

The combination of sets by configurations resulted in statistically significant effects indicating that certain combinations of figure-ground orientation and configurations were producing more rotation than others.

REFERENCES

- Asch, S. E. & Witkin, H. A. Studies in space orientation. II, Perception of the upright with displaced visual fields. *J. exper. Psychol.*, 1948, 38, 325-337.
- Bahrack, H. P. Incidental learning under two incentive conditions. *J. exper. Psychol.*, 1954, 47, 170-172.
- Bakay, E. & Schiller, P. H. Manipulative correction of visually presented figures. *Amer. J. Psychol.*, 1948, 61, 487-502.

- Baroff, George S. Bender visual-motor function in mental deficiency. *Amer. J. ment. Defic.*, 1957, 61, 753-758.
- Bender, Lauretta. A visual motor gestalt test and its clinical use. *Amer. Orthopsychiat. Assoc. Research Monogr.*, 1938, No. 3.
- Bender, Lauretta. *Child psychiatric techniques*. Springfield, Ill.: Charles C. Thomas, 1952.
- Bensberg, G. J. Performance of brain-injured and familial mental defectives on the Bender Gestalt. *J. consult. Psychol.*, 1952, 16, 61-64.
- Boring, Edwin G. Visual perceptions as invariance. *Psychol. Rev.*, 1952, 59, 141-148.
- Born, M. A. & Rubenstein, H. Information theory and immediate recall. *J. exper. Psychol.*, 1952, 44, 260-266.
- Broadhurst, P. L. The interaction of task difficulty and maturation: the Yerkes-Dodson law revived. *Acta Psychologica*, 1959, 16, 321-335.
- Byrd, E. The clinical validity of the Bender-Gestalt test with children. *J. proj. Tech.*, 1956, 20, 127-136.
- Casperon, R. The visual discrimination of geometric forms. *J. exper. Psychol.*, 1950, 40, 668-682.
- Clawson, Aileen. The Bender visual motor test as an index of emotional disturbance. *J. proj. Tech.*, 1959, 23, 198-206.
- Duffy, E. Leeper's motivational theory of emotion. *Psychol. Rev.*, 1948, 55, 324-328.
- Easterbrook, J. A. The effect of emotion on cue utilization and the organization of behavior. *Psychol. Rev.* 1959, 66, 183.
- Fabian, A. A. Vertical rotation in visual motor performance, its relation to reading reversals. *J. educ. Psychol.*, 1945, 36, 129-154.
- Feldman, I. S. Psychological differences among moron and borderline mental defectives. *Amer. J. ment. Defic.*, 1953, 57, 484-494.
- Fuller, J. B. Factors influencing rotation in the Bender-Gestalt performance of children. Unpublished doctoral dissertation, Univer. of Ottawa, 1960.
- Galloway, E. & Dembe, E. Narrowed attention: A psychological phenomenon that accompanies a certain physiological change. *Arch. Neurol. Psychiat.*, 1958, 79, 74-90.
- Goldberg, F. H. The performance of schizophrenic retarded and normal children on the Bender-Gestalt. *Amer. J. ment. Defic.*, 1957, 61, 548-555.
- Goldstein, K. & Scheerer, M. Abstract and concrete behavior. An experimental study with special tests. *Psychol. Monogr.*, 1941, 53, 32-57.
- Hanvik, L. J. Note on rotation in Bender-Gestalt test. *J. clin. Psychol.*, 1953, 9, 399.
- Helson, H. & Fehrer, E. V. The role of form in perception. *Amer. J. Psychol.*, 1932, 44, 79-102.
- Hick, W. E. On the rate of gain information. *Quart. J. exper. Psychol.*, 1952, 4, 11-26.
- Koppitz, Elizabeth. The Bender-Gestalt test and learning disturbances in young children. *J. clin. Psychol.*, 1958, 14, 292-295.
- McPherson, Marian & Pepin, Loretta A. Consistency of reproductions of Bender-Gestalt designs. *J. clin. Psychol.*, 1955, 11, 163-166.
- Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
- Shapiro, M. B. Experimental studies of a perceptual anomaly. II. Confirmatory and explanatory. *J. ment. Sci.*, 1952, 98, 605-617.
- Williams, Harold L., et al. The relation of brain injury and visual perception to block design rotation. *J. consult. Psychol.*, 1956, 20, 275-280.
- Witkin, H. A. & Asch, S. E. Studies in space orientation. III. Perception of the upright in the absence of a visual field. *J. exper. Psychol.*, 1948, 38, 603-614.
- Yates, A. J. Experimental studies of a perceptual anomaly. IV. The effect of monocular vision on rotation. *J. ment. Sci.*, 1954, 100, 975-979.
- Yates, A. J. Rotation of drawings by brain-damaged patients. *J. abnorm. soc. Psychol.*, 1956, 53, 178-181.

Received May 8, 1961

Responses of Children to a Group Projective Film and to the Rorschach, CAT, Despert Fables and D-A-P

MARY R. HAWORTH
Nebraska Psychiatric Institute

Preliminary investigations with *Rock-A-Bye, Baby*, a group projective film for children, suggest it may be a suitable instrument for use as a screening device for detecting adjustment difficulties in school populations. The film has been previously described (Haworth, 1957, 1960; Haworth and Woltmann, 1959). A repeat study (Haworth, 1961) reported close correspondence between proportions of deviant responses secured from two independent samples of approximately 250 children each. The developmental progressions observed in the responses of 500 children from kindergarten through fifth grade suggest that dynamic factors are being tapped which are pertinent to the Oedipal and early latency periods.

The present study was designed to assess the construct validity of responses to the film. A battery of individual projective techniques (Rorschach, CAT, Despert Fables, D-A-P) was administered to children who received the highest scores (most deviant) on certain indices of the film test and to a matched control group who did not score high on any index. The projective protocols were then to be evaluated for evidences of emotional difficulties along the same dimensions for which the experimental group was originally selected.

To the author's knowledge only one other study has been concerned with the consistency with which Ss react to a film and to an individual projective, in this case the Rorschach. Spiegelman (1955) hypothesized that differences in the perception of a film are a function of global aspects of personality. After showing a film depicting humans in action, he asked 24 adult Ss open-ended questions concerning the

activity and characters. Each subject was also administered the Rorschach. Judges were able to match film and Rorschach protocols significantly better than chance.

The present study was designed in an attempt to answer the following questions:

- 1) To what extent can judges differentiate experimental and control Ss on the basis of the individual projective test battery?
- 2) Will the groups differ in specific kinds of responses to the individual tests?
- 3) Within the experimental group, can the personality dimensions evidenced in the film responses (and used as the basis for selection) be observed in responses to the individual projectives?

Admittedly, the procedure has several shortcomings common to most attempts to validate projective findings: pitting a new projective test against other projectives also lacking in well-established outside criteria of validity; using clinical judgments as one of the bases for sorting; and employing "blind" and group analyses of children's protocols. It should also be pointed out that both groups of children are drawn from a presumably "normal" school population and none had been referred to a clinic.

METHOD

In the repeat study mentioned earlier (Haworth, 1961), the film test was administered to 257 children in kindergarten, first and second grades from two different schools. The objective scoring scheme designates deviant (atypical) responses in the following areas: identification, jealousy, aggres-

sion to parents, guilt (masturbation), anxiety (castration), and obsessive trends. The last three dimensions were selected as being, clinically, the most useful areas tapped by the film and those which prior testing (Haworth, 1957, 1961) had demonstrated most clearly reflect the child's manner of adjustment to the Oedipal situation.

Subjects

When the film test was administered, 23 children scored high on the Obsessive Index or on *both* the Guilt and Anxiety Indices.¹ When individual testing was begun the following year, only 15 of these children were still enrolled in their respective schools. Fifteen controls, matched for sex, school and grade at the time of the film test, were selected from the pool of 58 children who had received no deviant scores on any of the six film indices.

The final experimental group consisted of:

High Guilt and Anxiety (G-A):

4 girls, 5 boys

High Obsessive (OBS):

3 girls, 2 boys

High G-A and OBS:

1 girl

In all subsequent analyses, the one child who scored high on all three measures will be included with the obsessive group since she received the highest score of any child in this area. This gives nine G-A and six OBS cases. Grade distribution of the total sample at the time of individual testing was as follows: nine in first grade,² 17 in second grade, four in third grade.

¹ There were other children in the school sample who scored high on *either* Guilt or Anxiety, but not on the two in combination. Four of the six Obsessive cases also scored high on Guilt. It was also possible for children in the experimental group to have scored high on an index not under study.

² One of the experimental cases was repeating first grade. He was matched with a second-grader who had been in his room at the time of the film test.

Procedure

Since the author had administered the film test, served as one of the scorers, and made the final selection of experimental and control cases, she was familiar with the names of the children in both experimental and control groups. To insure against bias, these names (unaccompanied by group designations) were given to another examiner³ who administered the individual projectives to each child in the following sequence: Rorschach, CAT, Despert Fables. Each child was also asked to recall the puppet film. The Draw-A-Person Test was administered to the entire classrooms in which experimental (E) or control (C) children were enrolled. There was an interval from eight to twelve months, for each child, between the initial film test and the individual battery.

The individual examiner assigned code numbers to each of the 30 sets of protocols, so all judges were unaware of S's classification when making their selections. Each judge was instructed to select the 15 cases which showed the most disturbance in either the Obsessive or the Guilt-Anxiety dimension. No effort was made to select the correct number of OBS and G-A cases; in fact only one judge was aware of the exact breakdown.

It was necessary to set up a common frame of reference by defining the specific areas which the film indices presumably were measuring. The G-A film responses appear to reflect guilt associated with forbidden acts (mostly masturbatory) and fears of retaliation and bodily harm (e.g., castration). Apparently the conflicts and anxieties have been somewhat internalized, but the Oedipal situation has not as yet been satisfactorily resolved. Thus the Guilt-Anxiety dimension is viewed as the equivalent of anxiety-phobia, or of anxiety-hysteria, which Fenichel

³ Dr. Mary Jane Keller administered the individual tests and served as one of the judges. Appreciation is extended to her and to Gilbert DeRath who also served as a judge.

(1945) sees as the commonest form of neurosis in childhood.

The Obsessive dimension was defined in terms of the usual clinical orientation: severe super-ego with rigid defenses, such as over-meticulousness, intellectualization, isolation, reaction formation, etc. Fenichel (1945) states: "The great importance of the Oedipus complex, of castration anxiety, and of masturbation in the typical compulsion neurosis is well established" (p. 306). He further points out that compulsive rituals are apt to appear, along with the development of intellectual faculties, during latency.

In the children within the age range under study (5-8 years) difficulties encountered in the resolution of Oedipal problems are responsible for both types of neurotic mechanisms being examined. The defenses used to bind the anxiety differ in nature and extent, with repression featuring prominently in hysteria while reaction formation, undoing and isolation are characteristic of compulsion neuroses (Fenichel, 1945).

Three judges evaluated the protocols. In the first "round", Judges A and B were free to use any method of selection they wished to devise, in other words to rely on some form of clinical judgment.

Judge A reported being guided by the following criteria in making clinical judgments: marked blocking or rejection of cards, confused identification, perseverated themes, indications of guilt (masturbatory fantasies, aggressive impulses), unmet dependency needs, and certain Rorschach indicators (bizarre, minus, contaminated or confabulated percepts, color naming, numerous *m*, *K*, *C'*, *c*, and pure *C*.) Also considered were stories ending in death or destruction for the child character and employing any of the following themes: kidnapping, deprivation or abandonment, assault, aggression and violence.

Judge B (the author) developed

certain indices which she used as the basis for her final choices. These were, in turn, given to a third judge (C) to determine the extent of agreement between two judges when using the same objective criteria for their selections. Differences between experimental and control groups on the four tests were then examined, as well as any differences in the responses of the OBS and G-A groups.

RESULTS AND DISCUSSION

Agreement Between Judges

Using their preferred methods of arriving at clinical judgments, Judge A correctly identified nine of the 15 experimental cases, which was not significantly greater than chance, while Judge B selected 14 of the 15 experimental cases ($p = < .005$, Fisher's exact probability). The two judges agreed on eight of their selections (non-significant).

Judge B had set up, for each of the four projective tests, scoring criteria based on types of responses which seemed likely to be relevant to the dynamics under consideration. cursory inspection of the protocols eliminated some of the initial items, e.g., if 20-25 of the 30 children gave a certain response it obviously could not prove to be discriminatory. (The items finally selected will be discussed in more detail in the next section). Applying the scoring criteria, the protocols for each test were divided as nearly as possible into equal groups of presumed deviant and non-deviant, resulting in 15 deviant Rorschachs, 15 CATs, 17 D-A-Ps and 16 Fables. The final task of selecting 15 probably experimental cases involved first choosing the eight cases scoring on all four measures; then the four cases high on Rorschach, CAT and one of the other tests; and one case high on only the Rorschach and CAT. Qualitative inspection was used to select the final two cases from three cases which scored high on either Rorschach or CAT in combination with one of the other

two tests. As previously stated, fourteen of the fifteen cases so chosen proved to be correctly designated.

Since inspection of Judge B's results indicated that the criteria developed for the D-A-P and Fables would not have selected better than chance and that both the Rorschach and CAT criteria were highly successful, only the protocols of the latter two tests were given to Judge C, along with the scoring criteria and the definitions of the OBS and G-A dimensions for use in the final selection. Judge C was able to select 12 of the 15 experimental cases ($p = < .005$, Fisher's exact probability).

Judge C agreed with Judge B on the selection of 12 experimental cases

as disturbed and on 12 control cases as not disturbed. In addition, the two judges agreed on their mis-selection of one E case as non-disturbed and one C case as disturbed. They did not agree with each other in the placement of four cases.

All three judges agreed on seven E cases, two judges agreed on six more E cases, while two E cases were selected by only one judge. Only two C cases were judged disturbed by agreement of two judges. Table I tabulates these results as well as the extent of agreement between Judges B and C on the Rorschach and CAT. Using the objective criteria, the two judges agreed on 11 of the E cases for the Rorschach and mis-selected one control; they

TABLE I. Selections by Judges of Deviant Cases on Global Evaluation, Rorschach and CAT

Film Cases		Global			Rorschach		C A T	
		Judge A	Judge B	Judge C	Judge B	Judge C	Judge B	Judge C
OBS	1	+	+	+	+	+	+	+
	2	+	+	+	+	+	+	+
	3	+	+	+	+	+	+	+
	4	+	+	+	+	+	+	+
	5	+	+	+	+	+	+	+
	6	+	+	+	+	+	+	+
G-A	1	+	+	+	+	+	+	+
	2	+	+	+	+	+	+	+
	3	+	+		+		+	
	4	+					+	
	5		+	+	+		+	+
	6		+	+	+		+	+
	7		+	+	+	+	+	+
	8		+	+	+	+	+	+
	9		+	+	+	+	+	+
Totals		9	14	12	14	11	13	11
Controls	1	+						
	2	+					+	+
	3	+						
	4	+		+				+
	5	+						
	6	+						
	7						+	+
	8		+	+	+	+		
	9			+				+
	10					+		
	11							
	12							
	13							
	14							
	15							
Totals		6	1	3	1	2	2	4

TABLE II. Frequency of Certain Rorschach Patterns and Signs

	OBS (N=6)	Experimental G-A (N=9)	Total (N=15)	Control Total (N=15)	P
Obsessive pattern	5	0	5	0	0.25
Ames danger signals	1	8	9	1	0.05
Castration symbols (2 or more items)	2	5	7	0	
F+ % (70 and below)	4	6	10	2	
Smoke	0	0	0	5	
C + $\Sigma C > 3$	1	4	5	1	

agreed on 11 of the cases for the CAT and mis-selected two controls.

It should also be pointed out that of the six cases mis-selected by Judge A in the global evaluation, three were scored high on the CAT by one or both of the other judges, indicating that these were at the least equivocal cases.

Objective Clinical Criteria

Each of the tests used will be discussed in a separate section. Except where otherwise indicated the procedures and findings are based on those of Judge B. In the tables accompanying each section significance levels (using Fisher's exact probability method) are indicated where prior predictions were made. The remaining comparisons were made *post facto* but are included as possibly fruitful areas for further research with larger samples.

1. Rorschach

The protocols were scored by the Klopfer method, supplemented with Beck's criteria for F+. Each child's response summary was recorded on a master sheet, which also included possible castrative and masturbatory symbols and categories designed to reflect obsessive characteristics. The latter were grouped under the following headings: a) ambivalence, alternatives and doubting; b) expressions of balance and symmetry; c) overly-critical and qualifying statements; d) expansive, impulsive responses on Cards IX or X.⁴ The first three items were later combined and will be referred to as "a-b-c," and the expansive responses as "exp." In the search for meaningful

combinations of determinants, the following pattern was observed in five protocols:

1. M, 2 or more
2. ΣC , 3 or more
3. FC, 2 or more responses
4. a-b-c, 5 or more responses
5. Either, low F+ % (70 or below)
or, high F+ % (90 or above) plus exp

The five cases which met all of the above criteria were selected as the most likely to be obsessive.

All 30 protocols were next checked for the presence of the revised "Ames danger signals." Ames (1959) found the following items distinguished between disturbed and normal child Ss:

1. Excessive manipulation.
2. F% exceeds F+ %.
3. One or more C or Cn.
4. 70% A or more.
5. No H, or Hd:H is 2:1.
6. Static perseveration.
7. Contaminations.
8. Confabulations.
9. Position responses.
10. Elimination or sex response.
11. Bizarre content; personal reference; troubled content.

On an *a priori* basis it was felt these signs might be tapping manifestations of extreme guilt and/or anxiety. Using a score of 3 as a cut-off point (as recommended by Ames) 10 cases were selected by this measure.

Fifteen cases were thus selected by Judge B using these two analyses and 14 of them subsequently proved to be

⁴ This feature of obsessive protocols is suggested by Schafer (1954, p. 53).

experimental subjects ($p = < .005$). As can be seen in Table II, the obsessive pattern did, in fact, select five of the six OBS cases and no controls ($p = .025$) while the Ames signs selected eight of the nine G-A cases, the one remaining OBS, and mis-selected one control ($p = .005$).

Using the same process of analysis, Judge C picked 13 cases as belonging to the E group and 11 of these were correct ($p = .025$). He agreed with Judge B on the five cases selected with the obsessive pattern, on six G-A cases selected by the Ames signs and on one case mis-selected by the Ames signs, (see Table I).

Obsessive pattern. Remarkably close agreement (5 out of 6 cases) was obtained with the obsessive criteria. Four of the items (*M*, *FC*, ΣC and $F + \%$) are quite objective. Difficulty was anticipated in scoring the *a-b-c* category, but the judges agreed on 11 cases, of which nine were in the experimental group (four G-A's in addition to five OBS). Two additional E cases and one C child were selected by only one judge. The widespread use of the *a-b-c* items in the E group, as contrasted to the controls, suggests that the thought processes represented by this cluster are not solely the correlate of obsessive mechanisms but are indicative of the presence of neurotic difficulties in general.

The obsessive pattern developed in this study resembles in many respects the configuration found by Krugman (1946) in stuttering children and which he mentions also observing in obsessive children. Aside from Krugman's brief reference, very little has been formally written concerning obsessive Rorschach patterns in children or the extent of similarities between adult and child records. Certainly, initial reading of the protocols in the present sample revealed very little in the way of over-verbalizing, intellectualizing or attention to minute details. Yet such features were prominent in certain CAT protocols which were

subsequently found to have been given by the obsessive group.

Aside from the absence of small detail responses, the chief difference between the present obsessive pattern proposed for children and Schafer's (1954) criteria for adults consists in a higher sum *C* for the child group. Schafer notes that if the obsessive adult's defenses begin to fail, then an increase in color responses and a decrease in $F + \%$ become evident.

Both Schafer (1954), with adults, and Halpern (1953), with children, note the presence of sudden impulsive, explosive responses in the records of obsessive-compulsive subjects. Schafer (1954) finds such "sloppy" responses occur frequently on cards IX or X (in adults) and regards them as indications of poor integration.

Several references to the obsessive-compulsive syndrome in childhood (Bellak, 1954; Fenichel, 1945; Freud, 1959; Halpern, 1953; Lindemann & Dawes, 1952) suggest that initial indications may be observed in behavior and in projective responses within the age range of the present sample. The high sum *C* and expansive reactions noted in the Rorschachs probably represent an obsessive condition "in process" before defenses have become solidified, rather than signifying the crumbling of controls as would be the case with adults. It is suggested that what is being observed in the present sample is an incipient obsessive character structure. If and when defenses become more rigid, then color responses may be reduced, more attention may be paid to small details, over-verbalization and intellectualization may occur, etc.

Ames danger signals. Using the Ames signs Judge B was successful in selecting eight of the nine G-A cases (Table II), while Judge C selected six of these same eight cases (see also Table I). The lack of clear-cut agreement reflects the need both judges felt for more precise definitions for several of the items. Nevertheless, general

agreement was noted in the mean scores of each judge and as compared to those reported by Ames (1959). For a disturbed group of 50 boys (ages 6-12) Ames reports a mean 3.1 as opposed to a mean score of 0.6 in her normal control group. In the present study the respective mean scores for the two judges were: E group = 2.73 and 2.46; C group = 1.4 and 1.27. Further breakdown of the E group shows: OBS = 1.5 and 1.8; G-A = 3.55 and 3.11.

Ames (1959) points out that normal seven-year-olds give more disturbed signs than any other age. Over half of the present sample was composed of seven-year-olds; nevertheless discrimination was still possible between the E and C groups. In the present study, the Ames signs appear to be most effective in selecting anxious, fearful, guilt-ridden children but do not pinpoint children who seem to be in the process of developing obsessive characteristics. Two factors may be responsible: none of the children in the present sample was disturbed enough for clinic referral; and the age range was more restricted than that of Ames. Consequently, it seems highly possible that as an obsessively oriented child progresses through latency, he may subsequently score high on the Ames signs in addition to giving "obsessive" responses.

Castration symbols. The response items selected as possibly indicative of castration anxieties did not occur frequently enough to warrant using them as a basis for selection. However, *post hoc* inspection and comparisons between groups suggest these four castration categories which might comprise an index:

1. Tail, horns or feet missing
2. Parts cut or torn off; something eaten up
3. "Dead" responses
4. Blood responses

Seven experimental children (five G-A and the OBS girl who also scored high on G-A) gave two or more re-

sponses in these categories while none of the controls did so (Table II). This lends support to previous dynamic formulations that fears of injury are reflected in the Anxiety Index of the film test.

The content categories selected as possible masturbatory symbols occurred too frequently to discriminate between sexes or groups.

Additional diagnostic signs. After the selections were made other miscellaneous Rorschach responses were inspected for possible discriminatory usefulness (see Table II). $F+%$ alone would have selected fairly well, with 10 Es scoring 70 or below as compared to only two controls. The frequency distribution is even more striking (Table III). The average $F+%$ of the E group is 66.4 and of the C group, 82.0.

TABLE III. Frequency Distribution of $F+%$ in Experimental and Control Groups

	Experimental (N=15)	Control (N=15)
91-100	3	4
81-90	0	5
71-80	2	4
61-70	2	1
51-60	4	0
41-50	2	1
31-40	2	0

Responses of clouds, fire, *M-*, *FM-*, *m*, *S*, *Fc* were as apt to be given by C as by E cases. In contrast to expectations, smoke responses were given by one-third of the Cs and none of the Es. Although *C'* responses were given equally by both groups, the combination of *C'* with $\Sigma C > 3$ was given by five Es (four were G-A) and by only one control. This finding would be congruent with the anxiety-hysteria syndrome postulated for the G-A group.

2. CAT

All 30 protocols were initially examined for types of responses and test reactions which might be considered indicative of the presence of various defense mechanisms. In the com-

TABLE IV. Frequency of High Scores on CAT Defensive and Identification Measures

	Experimental		Control		P
	OBS (N=6)	G-A (N=9)	Total (N=15)	Total (N=15)	
1. Reaction, Undoing	5	4	9	1	.005
2. Isolation	6	5	11	2	<.005
3. Deception	3	4	7	4	ns
4. Denial, Repression	5	3	8	4	ns
5. Guilt, Fear, Anxiety	4	2	6	2	ns
6. Projection, Introjection	4	3	7	3	ns
7. Symbolism	2	4	6	1	.05
8. Regression	1	1	2	0	ns
9. Weak Controls	3	1	4	0	.05
10. Confused Identification	4	7	11	2	<.005

pleted schedule, six or eight sample descriptive responses were listed under each of the following defenses:⁵

1. Reaction - formation, undoing, ambivalences (e.g., oppositional attitudes, indecision by S or story character).

2. Isolation (e.g., detached attitude, "it couldn't happen").

3. Deception (e.g., adult tricks child).

4. Denial and repression (e.g., child character waits, is good, learned lesson).

5. Guilt, fear, anxiety (e.g., child character hides from danger).

6. Projection or introjection (e.g., attacker is attacked).

7. Use of symbolism (e.g., fire, explosions, rope breaks).

8. Regression (e.g., dirt and messes responses).

9. Controls weak or absent (e.g., perserverations, bizarre content).

Tabulations of all the protocols were made, and cut-off points were established between high and low "scores" for each defensive area. Insofar as the data permitted, scores falling in the upper two-thirds of a range were considered to be high. (E.g., the highest number of Isolation responses given by any child was 12, so all children giving five or more responses in

this area were considered to be high on Isolation).

A measure of deviant identification was devised by determining, whenever possible in the stories, if the child seemed to be identifying with a male or a female figure. If the sum of stories with opposite-sex identifications was equal to or greater than the sum of same-sex identifications, the child was considered to be "confused" as to identification.

The total number of high scores received by each child on the ten defensive and identification items was determined. The range was from 0 to 9. Fifteen children had high scores in three or more areas so these were considered by Judge B as nominations for the E group. Table I indicates that 13 of these cases were correctly designated ($p = < .005$). Using the same procedures Judge C found 15 children with three or more high scores, of which 11 were correct ($p = .025$). The two judges agreed on 11 E cases and on 2 mis-selected C cases.

In addition to the over-all high scores discussed above, it was predicted that more of the E than the C cases would have high scores on each individual defense and identification measure and that within the E group more of the presumably obsessive children would score high on the Reaction-Undoing and Isolation measures while the Guilt-Anxiety cases would have more high scores toward the Guilt, Projection and Regression end of the continuum.

⁵ Mimeographed copies of the CAT schedule in its tentative form may be secured from the author.

Table IV shows that on five of the ten measures significantly more E than C cases scored high, and that on the remaining five measures the trend was in favor of the E group. Confused identification was shown in 11 E cases as contrasted to two controls. Eight Es received high scores in five to nine areas, and five of these were OBS. No control had as many as five high scores. No significant differences were found within the E group but outstanding was the finding that all six OBS scored high on Isolation and five out of the six OBS scored high on Reaction-Undoing and on Denial-Repression.

While the agreement between the Judges B and C was significant with respect to their final choices of those cases with three or more high scores, the actual numerical scores assigned by each judge varied by two or more points in eight cases and by one point in 13 cases. Consequently, the scale needs further refinement before it can be used with confidence in individual evaluations.

The average number of high defensive scores for each sub-group showed a consistent trend between Judges B and C respectively: OBS, 6.06 and 4.17; G-A, 3.77 and 2.88; total E group, 4.73 and 3.40; total C group, 1.26 and 1.60. The magnitude of the group means is in the expected direction.

In the discussion of the Rorschach it was noted that none of the protocols showed obvious over-verbalization. Qualitative evaluation of the CAT protocols revealed six cases exhibiting signs of excessive verbalism and over-intellectualization. (This was reflected to some extent also in their high scores on Reaction-Undoing and

Isolation). Five of these cases proved to be in the OBS group.

Specific response patterns. Certain patterns seemed unique and were subsequently evaluated for dynamic significance (Table V).

Attitudes toward adults. While many children (17 out of 30) gave responses indicating the smaller or child figure tried to trick or fool the adult, only 11 employed the reverse tactic whereby the adult was viewed as plotting against the child and or not really being what he appeared to be on the surface. Seven of the 11 instances were in the E group; with four of the six OBS responding in this fashion (Table V). Dynamically, such a response suggests considerable distance between parent and child. It would not be surprising if such distrust would create difficulties in identification with a parent who appears to be "two-faced" and unreliable. Eight of these 11 children did also show confused identification in the CAT stories.

Attack themes on Cards 6 and 9. Analysis of stories told to the cave scene (Card 6) and the bunny alone in bed (Card 9) revealed distinctive patterns for the six E and eight C children who used themes of attack. The six E cases all told such stories to both cards, with stories to Card 6 usually expressing fear of possible attack and stories to Card 9 always describing actual attacks on the child character from wild animals, ghosts or hunters. Usually the attack in Card 9 was a fulfillment of the specific fear previously expressed on Card 6.

In contrast, the eight C children rarely responded with attack themes to Card 6 or to both cards. Usually

TABLE V. Frequency of Special Response Patterns on CAT

	OBS (N=6)	Experimental G-A (N=9)	Total (N=15)	Control Total (N=15)
Adult tricks child	4	3	7	4
Fears attack, #6				
and actual attack #9	2	4	6	1
Rope breaks, #3	1	2	3	0

only Card 9 involved this theme and the fantasy attack did not actually take place. Instead, the Card 9 story expressed either fear of an attack or a dream of an attack; often by a hunter, but never by an animal. In all but one case the child figure was reassured by realizing it was just a dream, or the wind, or "nothing at all," or he hid until the danger was past.

The castrative elements of actual attack stories on Card 9 would lead to expectation of their occurrence more frequently in the G-A group. This was somewhat the case with four of the nine G-As (and two of the six OBS) giving such stories (Table V).

Broken rope (Card 2). Another castration symbol suggested by Bellak (1954) concerns describing the rope on Card 2 as being broken. Only three of the 30 children gave this response and all were in the E group (two G-As and one OBS). Card 7, which Bellak also regards as eliciting castration themes, did not differentiate between groups with respect to stories of the monkey being attacked or eaten, and no S gave a tail-biting story.

3. Despert Fables

Selection on the basis of the Fables (Despert, 1946) alone would not have discriminated between the two groups. This may be a function of the method of analysis used rather than any weakness in the test itself. Responses were tabulated for specific types of answers to specific fables, e.g.: mentioning the same-sex parent or not mentioning any family member on the Separation Fable (No. 4); fear of attack in the Anxiety Fable (No. 5); damage to the trunk on the Elephant Fable (No. 6), etc. Personal references, refusals, perseverations, and unusually ambivalent or detailed responses were also considered. Responses to the Separation Fable and personal references appeared to be the most useful items, but neither one discriminated at a significant level.

Subsequent inspection revealed in-

teresting sex differences in the responses to the Elephant Fable which is considered by Despert (1946) to tap castration anxieties (Table VI). Ten of the entire boys' sample of 14 expressed concerns of injuring, dirtying or shrinking of the elephant, with all seven of the experimental boys responding in this manner. Six girls (out of 16) gave similar "castrated" responses, but five of these girls were control cases. In contrast, only two boys (both controls) said that the elephant would get bigger or prettier, while six girls gave this response, and five of them were in the experimental group.

TABLE VI. Sex Differences in Type of Response to the Despert Elephant Fable

	Experimental (N=7)	Control (N=7)
<i>Boys:</i>		
Damage	7	3
Growth	0	2
Other	0	2
<i>Girls:</i>	(N=8)	(N=8)
Damage	1	5
Growth	5	1
Other	2	2

Such sex differences can be interpreted according to psychoanalytic principles. Boys in general, as well as a significantly larger proportion of the experimental boys, seem to be expressing castration fears via an animal generally acknowledged as stimulating phallic fantasies. The reversed dichotomies for girls suggest that the control group's use of "damage" responses may indicate recognition and acceptance of anatomical differences, while the greater use of "growth" responses by the girls' experimental group may be reflecting the presence of considerable amounts of penis envy.

4. Draw-a-Person Test

Although the experimental group could not be selected on the basis of the combination of D-A-P items used, the following individual items (see

Table VII) may show some promise in clinical evaluations:

a) *Inconsistent handling of the two figures* (e.g., one stick figure and one in two dimensions, one of a pair drawn as a cartoon or caricature). Only experimental cases employed this device, thus discriminating at the .01 level.

b) *Level of sexual differentiation*, using Haworth's scale (Haworth & Normington, 1961). Level 1 involves little or no distinction between the two figures and was used more often by the experimental group. Level 2, requiring at least a skirt and longer hair on the female figure, normally should occur the most frequently in the age range of this study and was found significantly more often ($p = .025$) in the control group.

c) *Treatment of opposite-sex figure*. The experimental group was more prone to draw the opposite-sex figure either first or larger, while the control group more often drew opposite-sex figures both first and larger, in the instances where the own-sex figure was not given the usual preferential handling.

d) *Omission of facial features*. (Eyes, nose and/or mouth on one or both figures). When this occurred, such omissions appeared more often in the experimental group.

TABLE VII. Frequency of Certain D-A-P Items

	Experi- mental (N=15)	Control (N=15)	P
Inconsistent handling	6	0	.01
Sexual differentiation:			
Level 1	8	3	ns
Level 2	5	12	.025
Level 3	2	0	ns
Opposite-sex figures:			
First or larger	6	2	ns
First and larger	1	5	ns
Omit facial features	7	2	ns

It should be noted that the first three items above all involve comparative treatment of the two figures which suggests that the D-A-P may be most useful in revealing possible conflict or

confusion with respect to own and opposite sex.

It was originally planned to use all six items from the study (Fox, Davidson, Lighthall, Waite & Sarason, 1958) of figure drawings by test anxious children in the first through fourth grades. Four of these items (smiling, rigidity, arms down, humor) occurred too frequently to be discriminating. The shading item was retained but was not successful; the mutilation item was altered to consider omission of facial features (which proved useful) and omission of hands, feet or limbs (which did not discriminate between groups).

Anxiety Responses

In the evaluation of each of the individual tests, it has been pointed out that certain responses appeared to reflect fears of bodily injury or damage: use of at least two "castration" items on the Rorschach; fear of attack and its subsequent occurrence on two specific cards of the CAT; damage to the elephant in the Fables (for boys); omission of facial features on the D-A-P. Since the girls' responses on the Fable raise questions as to the ambiguity in interpretation of castration symbolism for females, the following analysis will be considered only with the boys: five G-A, two OBS, seven controls.

Summing the number of times damage responses were given on all four tests results in 14 for G-A, five for OBS, and only five for the controls. In terms of individual children, all five G-A gave such responses on two or more tests, one OBS did so and only one control.

It will be recalled that the G-A group was selected on the basis of scores on both the Guilt and the Anxiety Indices of the film test and that no control child received high scores in these areas. The items on the Anxiety Index of the film concern aggression by or toward the witch, and preoccupation with symbolism generally

associated with the witch figure. It was postulated that responses in this index suggest the presence of castration fears and the consequent anxiety. The fact that all the "high-anxiety" boys have responded to similar dynamic symbolism on the individual tests approximately a year later tends to confirm the theoretical assumptions made with respect to the Anxiety Index of the film. Unfortunately the present sample is too small to be definitive.

Recall of the Film

After completing the individual test battery with each child the examiner asked whether he could remember the puppet movie seen the year before. Only 11 children could recall; seven of these were Es (three G-A, four OBS). Items recalled could be grouped to correspond to the five main sequences of the film: a) initial scenes of the parents scolding and of Casper not wanting to mind; b) Casper is jealous of the baby and talks about getting rid of her; c) he calls the witch and she puts a spell on the milk; d) Casper beats the witch so that she removes the spell; e) parents return home, the baby is well, everything ends happily.

The OBS (three of four) were the *only* group to give memories of the initial parent-child scenes. Three OBS also recalled the jealous sequence, all four described the witch's spell, but only two reported on the beatings. Their concern is thus largely with the obedience aspects of the story. All three G-As omitted any reference to the initial parent scenes and to the jealousy aspects, but all three recalled the two witch sequences. These were the only memories for two G-As, the third also recalled the final scene. No more than two (out of four) controls responded in any one sequence, and the first scene was the only one not mentioned by any of them.

The exclusive concern of the G-As with the witch scene again confirms

the strong impression which that character and themes of aggression have made on this group. In contrast, the OBS were the only group to concentrate in areas concerned with parents, obedience and naughtiness. The strong superego aspects of these responses were unique in another respect. In five of six instances where Ss were recalling punitive actions of the father or mother, they also were unable to recall the details accurately. These same blocked memories elicited spontaneous moralistic remarks, such as, "If he wanted to play jokes he should play them on himself." It should also be pointed out that these OBS remembered other sequences with remarkable accuracy.

The samples are small, yet the following observations seem in order: one-half of the Es and only one-fourth of the controls are able to recall any details of the film after almost a year's lapse of time; the G-As are exclusively preoccupied with the character and scenes which had contributed to their high scores on the Anxiety Index of the film test; the OBS reflect severe superego in their recall of parental punitiveness and in their confusion of details in these scenes as contrasted to their quite accurate and detailed memories for other sequences.

Conclusions

It has been shown that judges can achieve highly significant agreement using objective criteria for the evaluation of individual projective material and that these criteria can be used to differentiate between the children who scored high and low on a group test.

The results of the present study suggest that children do tend to give consistent responses, reflecting the potency of dynamic dimensions, irrespective of the projective medium employed. The dynamic origins of these responses are highlighted by the fact that such consistencies are also observable over the time span of 8-12 months between the original projections to the film and

the subsequent responses to the individual test battery.

The children judged to be obsessive on the film test were also found to score in uniformly similar fashion on certain presumably "obsessive" indices of the Rorschach and the CAT, as well as recalling aspects of obedience from the film. The outstanding verbal characteristics of this syndrome were more noticeable on the film and CAT than on the Rorschach.

While it proved difficult to pinpoint specific "guilt" responses on the individual projectives, the validity of the film's Guilt Index has been established against empirical criteria in an earlier study (Haworth, 1961). Anxiety responses were clearly evident in the G-A group on the individual tests, in terms of concern with bodily injury or damage, and on recall of the specific witch scenes of the film.

As stated previously the subjects were all part of the "normal" school population. The judges were impressed with the deviant nature of many of the sets of protocols, some of which would rival, by clinical standards, those of the most serious of referrals. It would appear that this particular age range marks the high point for "danger signals" as pointed out by Ames and her co-workers (1952, 1959). Similar findings were noted by Haworth (1961) in a study of developmental progressions in types of responses to the film test. She found a definite decrease between the second and third grades, in the proportions of children scoring high on three of the film indices: Aggression, Guilt, Anxiety. It may well be that we still have much to learn with respect to the range of "normality" in the projective responses of early latency children. It is also possible that the most deviant material secured at these ages may be predictive of later difficulties in adjustment which may subsequently necessitate referral and treatment.

The film test appears to select, as well as the individual projectives,

those children who are experiencing special difficulties with Oedipal problems, which are in turn indicative of early neurotic disturbances. The test can be administered to groups of 12-15 children at one time and employs an objective scoring scheme for initial appraisal. Less personnel time is required for administration, scoring and evaluation than would be needed for the administration alone of any one individual test to as many children. Consequently, the film test seems to have promise as a practicable and valid substitute for individual testing, especially where it is desired to secure an overall assessment of large numbers of children.

SUMMARY

This study was concerned with the extent to which a group projective film test can obtain personality information comparable to that secured from individual tests. Haworth and Wolmann's *Rock-A-Bye, Baby* was presented to 257 children in kindergarten, first and second grades. The 15 Ss responding most deviantly on either the Obsessive measure or a combination of Guilt and Anxiety were matched with 15 children who had not scored in a deviant manner on any of the six indices of the film. Eight to 12 months after the film test each child was administered the Rorschach, CAT, Despert Fables, D-A-P, and asked to recall the story of the film.

Judges evaluated all 30 sets of protocols for evidences of disturbance with particular reference to the same dimensions for which the experimental group was chosen. Using "clinical judgment" two judges could not discriminate between groups any better than chance. The objective criteria used by one judge for the Rorschach and CAT were given to a third judge resulting in highly significant agreement in selection by this method.

Consideration was also given to the type and extent of differences in responses of the experimental and control groups and, within the experi-

mental group, the degree to which the Obsessive and Guilt-Anxiety cases responded in keeping with the dynamics to be expected of their respective syndromes. For the Rorschach, items of an obsessive index devised for this study were given only by the experimental group (and exclusively in the obsessive sub-group). The obsessive index combined with the Ames danger signals significantly discriminated between experimental and control groups. A schedule of defensive and identification items was developed for the CAT and also successfully separated the two groups. While no overall measure proved successful with either the D-A-P or Fables certain of the individuals analyses of the D-A-P did discriminate.

Within the experimental group, the Obsessives were the only cases to score on the obsessive measure of the Rorschach. They also scored consistent with expectations on the CAT defenses and recalled punitive scenes from the film. High scores on the Ames danger signals were found predominantly in the Guilt-Anxiety group; specific responses relating to fears of bodily injury or castration were prominent for this group on the individual tests and in the witch scenes recalled from the film.

It would appear that children do tend to respond in dynamically consistent fashion on individual and group tests and over extensive periods of time. Since evaluations comparable to those secured from individual projectives can be obtained in a group situation with the film test, the latter instrument has demonstrated its usefulness as a quick screening device in the school situation.

REFERENCES

Ames, Louise B. Further check on the diag-

- nostic validity of the Ames danger signals. *J. proj. Tech.*, 1959 23, 291-298.
- Ames, Louise B., Learned, Janet, Metraux, Ruth W., & Walker, R. N. *Child Rorschach responses*. New York: Paul B. Hoeber, 1952.
- Bellak, L. *The TAT and CAT in clinical use*. New York: Grune & Stratton, 1954.
- Despert, J. Louise. Psychosomatic study of fifty stuttering children. I. Social, physical and psychiatric findings. *Amer. J. Orthopsychiat.*, 1946, 16, 100-113.
- Fenichel, O. *The psychoanalytic theory of neurosis*. New York: W. W. Norton, 1945.
- Fox, Cynthia, Davidson, K., Lighthall, F., Waite, R., & Sarason, S. B. Human figure drawings of high and low anxious children. *Child Developm.*, 1958, 29, 297-301.
- Freud, S. *Collected Papers*, Vol. II. (Transl. by Joan Riviere). New York: Basic Books, 1959.
- Halpern, Florence. *A clinical approach to children's Rorschachs*. New York: Grune & Stratton, 1953.
- Haworth, Mary R. The use of a filmed puppet show as a group projective technique for children. *Genet. Psychol. Monogr.*, 1957, 56, 257-296.
- Haworth, Mary R. Films as a group technique. In A. I. Rabin & Mary R. Haworth (Eds.), *Projective techniques with children*. New York: Grune & Stratton, 1960. Pp. 177-190.
- Haworth, Mary R. Repeat study with a projective film for children. *J. consult. Psychol.*, 1961, 25, 78-83.
- Haworth, Mary R. & Normington, Cheryl J. A Sexual Differentiation Scale for the D-A-P Test (for use with children). *J. proj. Tech.*, 1961, 25, 441-450.
- Haworth, Mary R. & Woltmann, A. G. *Rock-A-Bye, Baby: A group projective test for children*. (Manual and film). University Park, Pa.: Psychological Cinema Register, 1959.
- Krugman, M. Psychosomatic study of stuttering children: IV. Rorschach study. *Amer. J. Orthopsychiat.*, 1946, 16, 127-133.
- Lindemann, E., & Dawes, Lydia G. The use of psychoanalytic constructs in preventive psychiatry. *Psychoanal. Study Child*, 1952, 7, 429-448.
- Schafer, R. *The clinical application of psychological tests*. New York: Int. Univer. Press, 1948.
- Schafer, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- Spiegelman, M. Effect of personality on the perception of a motion picture. *J. proj. Tech.*, 1955, 19, 461-464.

Received May 4, 1961

Stylistic Consistency among Aesthetic Preferences¹

ROBERT H. KNAPP AND HELEN EHLINGER
Wesleyan University

About three decades ago Allport and Vernon (1933) published their classical study on expressive movement, which sought to demonstrate certain stylistic consistencies among expressive behaviors drawn from a number of different modalities. The present study in modest way seeks to pursue the essential logic of their work with respect to aesthetic preferences. We present here some evidence that just as varieties of expressive movement have been demonstrated to have stylistic consistency, so preferences for different artistic modalities reveal a stylistic congruence. If, as we believe, aesthetic preference is a subtle yet penetrating revelation of the temperamental and motivational attributes of the individual, there should indeed be some characteristic style common to an individual's aesthetic preferences. If, on the other hand, aesthetic preference proves to be specific to the particular modalities of artistic creation, then its importance, both theoretically and practically, for the evaluation of personality becomes questionable, or at least comprehensible only in much more complicated terms.

The present study constitutes a departure from the principle which has guided our previous research, in which we have sought to determine how personality qualities and characteristics are correlated with preferences for varieties of aesthetic material within a single modality, i.e., abstract art (Knapp & Green, 1960), tartan designs (Knapp, 1958; Knapp, Brimmer & White, 1959), poetic metaphors (Knapp, 1960). It will be realized that the present study does not seek to relate aesthetic preferences to

independent personality measures. Rather, it seeks to establish correlations among preference patterns for different aesthetic modalities.

METHOD AND PROCEDURE

In the present study we have administered three tests of aesthetic preference to 60 adult subjects, 32 men and 28 women, drawn from a population of teachers attending summer school at Wesleyan University. The first of these tests involved a measure of preference for Kodachrome slides of modern abstract art which has been reported elsewhere (Knapp & Green, 1960). This test consists of 25 such slides which have been factorially reduced to five groupings, and these five groupings, in turn, to two major groupings for which scores have been obtained in the present study. The first major group, Abstract Art A, combines Factors II and III of the Knapp and Green study (1960) and is characterized by a relatively high degree of geometrical patterning. Preference for this group has been shown to correlate negatively with neurotic tendencies, as manifested on the Minnesota Multiphasic Personality Inventory, and positively with certain extroverted interests and vocational predilections. The second major group of abstract paintings, Abstract Art B, is characterized by qualities of turbulence and the absence of clear geometrical principle. Preference for these has shown an opposite pattern of correlations with respect to the variables mentioned above. The method of scoring and administering this test has been reviewed in a previous article (Knapp & Green, 1960), and will not be repeated here.

The second test employed in the present study has not been reported heretofore. It consists of 16 short piano

¹ This research was supported by the National Institutes of Mental Health under Contract M-2178 to Wesleyan University.

TABLE I—Musical Preference Test

Melody No.	Tempo	Mode	Harmony	Selection
1.	Fast	Major	Dissonant	Debussy, <i>Pour Le Piano</i>
2.	Fast	Major	Consonant	Chopin, <i>A Flat Etude Opus Ten</i> #10
3.	Slow	Major	Consonant	Schumann, <i>Entreating Child</i>
4.	Slow	Minor	Dissonant	Debussy, <i>Perfumes of the Night</i>
5.	Fast	Minor	Consonant	Schumann, <i>Italian Sailor Song</i>
6.	Slow	Major	Consonant	Beethoven, <i>Fifth Piano Concerto in E Flat—Adagio</i>
7.	Fast	Minor	Dissonant	Debussy, <i>Pour Le Piano</i>
8.	Slow	Minor	Consonant	Chopin, <i>Prelude in B Minor</i>
9.	Slow	Minor	Dissonant	Rebikoff, <i>Nereid</i>
10.	Fast	Major	Consonant	Schumann, <i>Knight of the Hobby Horse</i>
11.	Fast	Major	Dissonant	Debussy, <i>Les Collines d'Anacapri</i>
12.	Slow	Minor	Consonant	Irish Folk Tune, <i>Foggy Dew</i>
13.	Fast	Minor	Consonant	Rachmaninoff, <i>Barcarolle G Minor</i>
14.	Slow	Major	Dissonant	Debussy, <i>Sunken Cathedral</i>
15.	Fast	Minor	Dissonant	Rebikoff, <i>Idyl</i>
16.	Slow	Major	Dissonant	Debussy, <i>Girl with the Golden Hair</i>

melodies from Romantic and Impressionistic compositions of the 19th century which were selected to incorporate permutations of consonance-dissonance, major and minor mode, and fast-slow tempo. (See Table I.)

These compositions were recorded on a tape, and this tape was played for the subjects twice, first in order to acquaint them with the variety and content of the melodies, and a second time to permit them to rate the melodies on a seven-point scale with respect to aesthetic appeal. The resulting ratings for the 60 subjects have been factor analyzed, and the factor loadings for each of the melodies on the two primary factors are reported in Table II. In scoring this test for the first factor, we have

summed the ratings assigned to melodies #4, #7, #9, #11, #13, and #15 and subtracted from this the rating assigned to melody #6. Of the six melodies with significant positive loadings, five are minor and dissonant in character, while melody #6, the selection from Beethoven's *Fifth Piano Concerto in E Flat*, is major-consonant in form. This factor is broad in nature and can best be described as representing melodies of a troubled, restless quality. The second factor shows significant positive loadings on items #12, #14, and #16 and negative loadings on items #1 and #5. The three melodies with positive loadings are all of slow tempo, while those with negative loadings are both of rapid tempo. The five selections do not differ with respect to the characteristics of harmony and mode. This second factor might be described as incorporating meditative, nostalgic melodies *vs.* animated melodies.

The third test which we have used in this study involves preference for schematic architectural designs. In this instance 12 linear drawings were prepared incorporating the permutations of the following variables: 1) cantilever, rectilinear, and curvilinear; 2) horizontal *vs.* vertical; 3) symmetrical *vs.* asymmetrical. The 12 architectural sketches were rank ordered by each subject in order of preference and scores obtained for the three pri-

TABLE II—Rotated Factor Loadings on Music Preference Test

Melody	Factor I	Factor III
1.	+ .01	— .27
2.	— .26	+ .07
3.	— .05	+ .19
4.	+ .50	+ .15
5.	+ .03	— .32
6.	— .42	+ .23
7.	+ .41	— .10
8.	+ .14	+ .03
9.	+ .50	+ .10
10.	— .16	+ .26
11.	+ .47	+ .08
12.	+ .05	+ .65
13.	+ .47	— .13
14.	— .11	+ .54
15.	+ .40	— .10
16.	+ .04	+ .44

mary modes as well as for the horizontal-vertical emphasis and the quality of symmetry-asymmetry. In the present study only the scores associated with preference for the three primary modes will be presented, since the differences in the horizontal-vertical and symmetry-asymmetry dimensions did not prove of statistical significance.

RESULTS

In Table III we present the matrix of intercorrelations among the seven variables involved in this study: the two musical scores, the two scores on the Abstract Art Test, and the three scores derived from the Architectural Preference Test. None of these correlations is corrected for attenuation, which, if done, would substantially increase their magnitude through their increased reliability. Nevertheless, a number of relationships stand out as significant. It will be observed that the correlation between Abstract Art A and Music II stands at $-.35$, well over the 1% level of confidence, indicating a secure negative correlation between preferences for the geometrical type of abstract art and for meditative, nostalgic music. It will be recalled that our Abstract Art A score involves a combination of Factors II and III, as reported in our earlier study (Knapp & Green, 1960): Factor II consists of rectilinear paintings, while Factor III are curvilinear, geo-

metrical paintings of Roger Kandinsky and Miro. Further examination of our data reveals the Factor II yields a negligible correlation in the present study with our second musical factor, but that Factor III yields a highly significant correlation of $-.41$. It is thus clear that this correlation primarily determines the negative relationship between preferences for curvilinear, geometrized art and meditative nostalgic music.

There is also a suggestive positive correlation between preference for turbulent and "neurotic" abstract art, Abstract Art B, and restless and troubled music, as represented in Music I.

The scores from the Architectural Preference Test yield an even more striking relationship to the Musical Preference scores, as shown in the matrix. Thus, preference for rectilinear architecture shows a significant negative correlation with Music I, and a suggestive positive correlation, just short of the 5% level of confidence, with preference for Music II. On the other hand, preference for curvilinear architecture (Architecture III) yields a secure positive correlation with preference for the first musical factor and an equally significant negative correlation with the second musical factor. Thus, it emerges that preference for restless, troubled music is positively associated with preference for curvilinear architectural forms and

TABLE III—Matrix of Intercorrelations Between Music Preference Test, Abstract Art Test and Architectural Preference Test.

	Music II	Abstract Art A	Abstract Art B	Architecture I	Architecture II	Architecture III
Music I (troubled, restless)	-.11	+.04	+.21	+.01	-.33	+.29
Music II (meditative, nostalgic)		-.35	+.07	+.16	+.23	-.33
Abstract Art A (geometric, rationalized)			-.22	-.22	-.16	+.32
Abstract Art B (turbulent, diffuse)				-.10	-.21	+.27
Architecture I (cantilever)					-.29	-.50
Architecture II (rectilinear)						-.68
Architecture III (curvilinear)						-

abstract art of a turbulent, diffuse character. On the other hand, preference for meditative, nostalgic music is associated with preference for architecture embodying rectilinear qualities and accompanies a dislike of abstract art involving curvilinear, geometrized forms. It may be noted that cantilever structure (Architecture I) yields no meaningful correlations with musical factors.

When the correlations between the abstract art scores and the architectural preference scores are examined, an interesting difference is revealed between the rectilinear and curvilinear architectural variables. Here both abstract art scores correlate negatively, though not significantly, with the rectilinear architectural score (Architecture II) and positively and significantly with the curvilinear score (Architecture III). It seems probable that this relationship is due to the fact that the ratings on the Architectural Preference Test fell into a forced distribution whereas those on the Abstract Art Test were made on a seven-point scale so that the mean of each subject's ratings varied about the rating scale mean of "4", although subjects were requested to try to use the full scale and to equalize their positive and negative ratings. It may be assumed that subjects not generally hospitable to abstract art tended to emphasize the negative side of the rating scale, while those favorably disposed tended to use the positive side. Such a rating bias would explain the correlation pattern reported here. This would suggest that a difference in modality preference is here revealed, so that those persons who are hospitable to abstract art generally tend to prefer the curvilinear architectural designs while those unreceptive to abstract art prefer the rectilinear type of architecture. One might say, at the risk of facetiousness, that the "squares" do not welcome the vanguard of modern abstract art.

Finally, it may be noted that the correlations among the three archi-

tectural scores are all negative and significant since the twelve architectural drawings were rank ordered in rating preferences so that the three primary scores bear a reciprocal and opposed relationship to each other.

SUMMARY

This study presents clear evidence for two different stylistic consistencies of preference across three modalities of aesthetic expression: architecture, music, and abstract painting. First, there seems to be a general stylistic consonance among (1) restless and troubled music, (2) turbulent and diffuse abstract art, and (3) curvilinear architectural forms. A second pattern of aesthetic preference combines (1) an interest in meditative and nostalgic music with (2) a dislike for curvilinear architectural forms and (3) a dislike for abstract art incorporating geometrical curvilinear designs. The dynamism of the first constellation with its neurotic overtones stands in contrast to the sober and calm qualities of the second. This distinction might be loosely identified with a stylistic dichotomy long recognized by critics of the arts. Thus, the first constellation bears some resemblance to the Dionysian and romantic traditions, and the second to the Apollonian and classical modes.

Beyond this, it should be noted that there is some evidence emerging of modality preference on the basis of the interpretation offered of the relations between the architectural scores and the abstract art scores. Here it would appear that persons preferring the more stable rectilinear structures tend to be less hospitable to the generally turbulent and chaotic character of modern abstract art. This study provides encouragement that continued research in this field will demonstrate deeper and broader consistencies among preferences for artistic styles in different modalities, thus confirming the hypothesis that aesthetic style may be deemed a general index to temperament and personality.

REFERENCES

- Allport, G. W. & Vernon, P. E. *Studies in expressive movement*. New York: Macmillan, 1933.
- Knapp, R. H. Achievement and aesthetic preference. In Atkinson, J. W. (Ed.) *Motives in fantasy, action and society*. Princeton, N.J.: Van Nostrand, 1958.
- Knapp, R. H. A study of the metaphor. *J. proj. techniques*, 1960, 24, 389-395.
- Knapp, R. H., Brimmer, Janet & White, Martha. Educational level, class status and aesthetic preference. *J. soc. Psychol.*, 1959, 50, 277-284.
- Knapp, R. H. and Green, S. Preferences for styles of abstract art and their personality correlates. *J. proj. techniques*, 1960, 24, 396-402.

Received July 3, 1961

The Utilization of the Psychologist-Patient Relationship in Diagnostic Testing

THEODORE LEVENTHAL

Worcester Youth Guidance Center

HOWARD J. SLEPIAN

Worcester Youth Guidance Center

MARTIN R. GLUCK¹

South Hills Child Guidance Center

BERNARD P. ROSENBLATT¹

Hampstead Child-Therapy Clinic

Within the past fifteen years, based on experimentation and clinical experience, growing attention has been focussed on the varied influences of the patient's role in the test situation. In the writings of people like Schachtel (1945), Sarason (1954), and Schafer (1954), the point is made strongly that responses are not solely a function of the test stimulus and the traits and abilities being measured, but also that such factors as set, attitudes, transference feelings, and resistances must be taken into account. While there is this recognition of the importance of the testee's or patient's perception of the situation, his intention, and his relationship with the examiner, the most common recommendation for testing purposes is that these data be used diagnostically, as well as to give a more accurate picture of the validity of the test information. As in traditional testing, however, the patient is allowed to structure the situation in whichever way he pleases. The tester in turn is seen as an *observer*; he intervenes in this structuring only when the testee refuses to conform to the test instructions, in which case something has to be done.

In contrast to procedures in which conscious and unconscious ideas and feelings are allowed free reign, and are either ignored or at best used diagnostically, this paper advocates an approach in which test-relevant thoughts and feelings are elicited and handled by the psychologist-examiner. The psychologist, by actively dealing with resistances, misconceptions, and fears, and by openly clarifying the

patient's needs as well as the test situation, strives to set up a testing relationship in which he is ideally perceived as a trusted *helper* rather than just an observer. While he cannot always attain this "ideal" relationship, he tries to move the testee in the direction of seeing the psychologist and his tools (tests) as being beneficial. The more this patient-helper relationship is established, the more will the patient's ego be enlisted in the diagnostic process.

Such handling of the test relationship has been found by us to have three broad consequences: 1) it has provided basically more valid and reliable test data, less influenced by uncontrolled situational factors and transference effects; 2) our understanding of both the relationship formed during testing and the context from which the test data has come have provided more meaningful diagnostic material; and 3) it has frequently made the testing experience psychologically advantageous and therapeutic for the patient.

In the next section literature relevant to the background of our view of the test relationship will be reviewed. In succeeding sections of this paper, an attempt will be made to describe in detail what is felt to be the most adequate testing situation, what is meant by "handling" the relationship, how such testing procedures are more helpful in diagnosis and how they are therapeutic. Finally, illustrative case abstracts will be presented.

BACKGROUND LITERATURE

In reviewing the literature about

¹ Formerly on the staff of the Worcester Youth Guidance Center.

the psychologist-patient relationship in diagnostic testing, it seems reasonable to approach the material in terms of the relative importance with which the relationship is regarded. The literature on testing technique may, on this basis, be arbitrarily divided into four categories. These are: 1) completely ignoring the test relationship, 2) recognizing the need for the patient's cooperation, or the establishment of "rapport", 3) recognizing the significance of the test relationship and its effect on test material and on the understanding of the patient (including experimental evidence as well as clinical contributions), and 4) recognizing the significance of the test relationship and utilizing it with the patient.

The review of the literature which follows will generally proceed in the order of the above categories. However, exact classification of many items is unclear, so that formal subdivisions into the categories listed is not attempted.

Clinical psychological testing in its early stages was concerned with measuring individual differences in various traits or abilities with a maximum of scientific objectivity. Toward this end, emphasis was placed on the "mental tests" themselves (e.g., Wells, 1927); for example, the almost exclusive attention paid to the standardization of administration and the standardization involved in establishing norms was aimed at objectifying testing procedures. The explicit intention was to present the subject with a situation in which both the test materials and the examiner's behavior were constants. The assumption was (and is) that if the test situation was presented in a standardized fashion, then the subject would perceive testing in a standard manner and all response variations would be solely a function of the subject's persistent traits or abilities.

But even where there was insistence on rigid, formalistic administration, its proponents frequently found it necessary to give some tacit or implied

recognition to the interpersonal nature of the examination. Thus, it was recognized that some of the variables which resulted in a test response could not be so simply controlled or standardized. It is in this connection that we encounter the much overworked and under defined term "rapport". For example, in describing the testing of children, Wells (1927) says, "The personality of the examiner is a governing factor in the attitude of the subject" (p.21). His suggestions for establishing rapport cover such manipulations as greeting a youngster with "Hello, what is your name?" (p.23), taking a walk with the child, and talking about the weather. His approach can be summed up in his statement that testing should be carried out with "kindness, firmness, and justice" (p.23). Some ten years later, very similar comments and directions are to be found in Terman and Merrill's Binet manual (1937). Here again there is emphasis on the need for rapport in order to obtain valid test results; the suggested methods for establishing this rapport revolve about friendliness, praise, sympathy, and the like. Similar discussions of the examiner-patient relationship can be found in most other reference works on testing (e.g. Wechsler, 1939, Klopfer and Kelley, 1942). Even recently in the area of vocational testing and counseling, very closely related definitions of rapport are to be found in an article by Sinick (1953). He too recognizes the effect of the examiner-tessee relationship on the test results, and makes suggestions which are much like those cited above.

The fact that the patient might not perceive the testing situation in a "standard" manner seems to have been regarded as an instance of poor rapport and thus as an interference with adequate testing. In fact, Bell (1948) makes this point explicit. In reference to projective techniques he states: "A second criticism is the sensitivity of the method to the interpersonal relationships between the ex-

aminer and the subject." (p.496). Here, then, the influence of interpersonal factors is viewed as a criticism and limitation of psychological tests. Bell goes on to point out this is a relatively unexplored area and that projective testing may possibly have variables in common with therapeutic techniques.

With increasing application of psychodynamically oriented concepts, there has been a growing recognition that not just the tests, but the personality characteristics (e.g. needs, defenses) of the examiner as well as of the patient may play a profound role in the patient's perception of and reaction to the whole process of being diagnostically tested. Ideas such as these have been given expression in the recent literature in two main channels. First of all, the implications of testing, the interpersonal situation it encompasses, and the results of this situation have been assessed from a primarily theoretical position; secondly, a good deal of experimental work in this area has been reported.

One of the foremost of the theoretically oriented papers is that of Schachtel (1945). Schachtel suggests that the psychoanalytic concept of transference may be useful in understanding a patient's reactions to being given a Rorschach. He points out clearly that patients may structure the situation in a variety of ways, dependent mainly on their own personality configurations. Schachtel explicates the parallel between the interaction in testing and that in interviewing. He does not, however, suggest that the examiner actively explore the patient's feelings and attitudes, or in any other way become active in the interpersonal situation. In point of fact, he says that the examiner becomes a "negligible factor" compared to the patient's set, except in "exceptional circumstances" (p.448).

The more recent of the systematic theoretical appraisals of the interpersonal relationship existing in the testing situation are represented in the

writings of Sarason (1954) and Schafer (1954,1956). Sarason elaborates on the many variables which are active in any face-to-face interviewing situation, and attempts to evaluate some of the possible effects of these variables in the psychodiagnostic setting. In his writing, Sarason stresses the necessity for the clinician to be cognizant of these variables, and to evaluate them. Apropos of this, he remarks that many psychologists "operate as if these variables were unimportant or did not exist" (p.85). All of this is aimed at making clinical judgment and test interpretation more accurate and comprehensive.

The general tenor of Schafer's *Psychoanalytic interpretation in Rorschach testing* (1954) is quite similar to Sarason's with at least one major difference in emphasis. Schafer explores more thoroughly the examiner's contribution to the test interaction. He dissects the relevant personality variables and pertinent sociological variables brought to the situation by the examiner and suggests how these, like their counterparts in the patient, may effect both individuals. As in all the approaches thus far reported, Schafer's writing stresses the need to be aware of these factors and their role in distorting or enriching the patient's productions and the examiner's clinical judgment. However, there is little said about structuring or dealing with the relationship, either for diagnostic or therapeutic purposes.

This is also largely true in Schafer's paper comparing the testing relationship to the transference situation in therapy (1956). After outlining the factors in the testing situation that foster transference feelings, he discusses the kinds of resistances patients are likely to use. There seems to be an implication that the psychologist's role in this situation should be 'negative'—that is, Schafer emphasizes that the tester does not actively elicit the charged material of a therapy session, nor does he confront the patient with

it as might be done in a therapeutic hour.

The second expression in the literature of this dynamically oriented assessment of the test situation is the experimental work which it has fostered. Research studies have varied from examining the effect of modifications of objective test instructions and administration, to studying the effects of different examiner personalities on projective protocols. Hutt (1947), in attempting to secure a more representative intelligence test score from Binet examinations, found that a departure from the standard instructions concerning the order of testing (i.e., alternating easy and hard items) seemed to aid in maintaining poorly adjusted children's motivation or attention. The relevant implication here is that by departing in a known way from the prescribed method, different kinds of information about the subject are more likely to be obtained. In a later paper, Hutt, Gibby, Milton and Patthurst (1950) attempted to vary the attitudes with which subjects viewed a Rorschach test and found that some Rorschach variables were affected by the attitudes induced in their subjects. As one of their conclusions, the authors state that the normal person is more likely to perceive and be affected by the situation in which he is placed, while the psychiatric patient may tend to operate in a more rigid way. This comment, similar in nature to Schachtel's would suggest that the examiner could do little to effect changes in a psychiatric patient's perception of the testing situation.

Lord (1950), using the Rorschach examination, found that there were quite consistent variations in types of records elicited by different examiners. These variations were consistent for each of the examiners, despite attempts on their part to play different roles with different subjects. Thus it is suggested that the examiner's own personality structure is a critical var-

iable in the testing situation. Gibby, Miller and Walker (1953), in a similar Rorschach study, also found significant variations in Rorschach records obtained by several examiners. They too follow Hutt's view by tempering their conclusions with the comment that more severely pathological patients (e.g., schizophrenics and obsessive-compulsives) may respond more in terms of their usual defenses rather than in terms of the current interpersonal relationship.

In addition to Hutt et al's (1950) study, there have been several others concerned with the effect of the environmental situation on test protocols. Klatskin's (1952) comparative study of the Rorschach protocols of women facing an emotionally significant situation and clerical workers on their regular job is representative of this group of studies. There was a significant difference in the patterns of the two sets of records.

More closely akin to Lord's approach is that of Luft (1953). Here the experimental variation in set was directly a function of the examiner's activity with the subject. He found that subjects who were treated in a friendly fashion by the examiner liked more of a set of cards they were asked to so judge than did the group of subjects treated in an unfriendly manner.

Viewing the problem from the point of view of "ego-involvement", Calden and Cohen (1953) also worked in the framework of experimentally varying the conditions under which the Rorschach test was presented to subjects. They too found similar results, and again qualify their findings by commenting that the less normal the subject, the less likely he is to be affected by such variables.

Although coming before Schafer's work chronologically, Cronbach's book on psychological testing (1949) seems to be in the middle ground between the more prevalent ideas on the examiner-testee relationship, and the

point of view of this paper. Although the significance of the examiner's role is not emphasized, there is much discussion of the testee's motivation and its effect on his test performance. Cronbach suggests that there is a *desirable* attitude of the patient toward testing, thus departing somewhat from the notion that all attitudes are equally desirable, provided they are known. He feels that the testee should want to take the tests, and that this cooperative attitude may be helped along by frankly explaining the reasons for the testing. (It is interesting to note that a very similar comment is made by Wells (1927) who stated that older subjects should be told the actual purpose of the examination in order to assure their cooperation. He concludes, "If the examination cannot be 'sold' on these terms, the patient is little able to profit by it" (p.23).) Cronbach further states that it is the job of the examiner to be aware of motivational and emotional factors, and to attempt to counteract or minimize them. For reasons not clearly presented, however, he says that this enlisting of the patient's cooperation is not a suitable procedure with children and clinical cases. With these, he suggests falling back on the examiner's "pleasing personality" (p.93).

Cronbach's comments on the uses of test results in a counseling situation, however, seem to be more forward. He remarks that the test results should not be for the psychologist's use alone, but that they should be used *with* the testee as an aid to his attaining understanding of himself. This approach, applied to clinical situations, is consistent with the views of the authors of this paper and has been expanded by limited numbers of research workers who have used such an approach with projective test materials. For example, Luborsky (1953), in an article on self-interpretation of TAT protocols, concludes that such a technique helps in evaluating the patient's self-concept, defenses, and prognosis. In a brief paper

challenging static concepts of the examiner-patient relationship, Joel (1949) also notes that modifications in testing methods may be needed, and that therapy and testing procedures may have variables in common.

Lastly, Holt (1956) has stated that "We cannot fully know . . . without associations and without knowledge of the subject what is condensed, what displaced, what symbolized in a Rorschach response . . ." (p.20). This view seems to us to be valid and to make a good case for a testing technique which will aid the subject in producing such associations.

In summarizing the work reviewed here, the following points appear salient: First, there has been growing recognition that the patient's perception of the testing situation and the examiner's behavior in it cannot be held constant; each of these varies from case to case. Theoretically, there are grounds upon which to compare the test-interaction with other interview situations, especially that of psychotherapy. There is experimental evidence to support the thesis that the test situation is vulnerable to many of the variables that effect any interpersonal situation. Finally, clinical writers are stressing the necessity for the practicing clinician to be aware of these facts and use them in his evaluations of the diagnostic picture. There are, however, few direct suggestions that the clinical psychologist might directly alter the situation, or deal with the individual personality and situational variables which the patient brings to bear on the examination. It is from this point that the present thesis proceeds.

THE "IDEAL" RELATIONSHIP

It is here proposed that psychological testing is most effective when there is an attempt to establish a particular kind of attitude and relationship on the part of the patient. Further, it is proposed that the ideal type of rela-

tionship in testing is one in which the testee looks upon the psychologist as a "helper" and then tries to work with the psychologist in the active process of understanding himself better. Although this "enlistment of the ego" of the testee in the diagnostic process cannot always be obtained, approximations of it, or any movement in this direction on the part of the patient, enhances the testing process and is beneficial in other ways as well.

Motivating the patient so that he *actively participates* in the testing process is obviously intricate and complicated. In the authors' experience it cannot be done through persuasion, or any devices which ignore the individual's dynamics (e.g., being uniformly re-assuring, friendly, etc.). The technique, rather, is based on interview and therapeutic principles. It is helpful, in fact, to look upon testing as being part of an interview, i.e., a face-to-face encounter in which there is a goal to be accomplished by both participants. The task to be accomplished is not that of completion of the tests, but rather that of obtaining a diagnostically accurate and thorough understanding of the patient by employing certain tools. Unfortunately, there has been an extreme centeredness on tests among psychologists, and even those who have recognized the limitations of such a narrow and instrument-oriented diagnostic approach have frequently been content to compromise by introducing a "pre-test interview". Although this can be useful, it all too frequently takes the form of asking certain stereotyped questions with the assumption that this is a prelude to the important work ahead—namely testing. The pre-test interview is generally of the question-and-answer type, without serious attempt to tap the testee's feelings.

The role of the psychologist should be perceived by the patient as that of a helper, rather than just an ob-

server. He attains this role to the extent that rapport, in Wolberg's (1954) sense, is established; i.e., the testee and his feelings, including misconceptions and resistances, are accepted and dealt with. It is only by exploring the attitudes and feelings of the patient and arriving at some clarity as to what the patient wants and needs, that the patient can be motivated to participate actively. The difference between test-centered diagnostic work and the approach put forth here may be likened to the difference between a fact-oriented, history-taking type of interview, and one in which the "process" or the relationship between interviewer and interviewee is emphasized. While the former is certainly easier to carry out, and, of course, supplies more factual information, what it misses frequently is *relevant* information, and moreover there is a greater risk of distortions of supposed facts. An important type of data—namely how the patient relates—is also less available. In addition, preparation of the patient for subsequent contact is absent. An interview that is process-oriented, however, does not delete factual information, but allows it to emerge in the context of a relationship in which one individual is trying to help another figure out what he wants, what his problem is, how he can be helped with it, and what may have caused it.

In order to help the patient relate in a realistic and affective manner, and have him feel that he is in an atmosphere of help and understanding, six overlapping areas are usually taken up as an essential part of the testing process as we carry it out. These are:

- 1) What is the problem as the patient sees it?
- 2) What does the patient see as the basis for the problem?
- 3) What are the patient's feelings about the problem—e.g., how concerned is he?

- 1) What can and should be done about the problem? What alternatives exist as possible approaches to solution?
- 5) How can testing play a role as a helping process, or part of one?
- 6) Interpretation of test findings to the patient in a way beneficial to him.

To help the patient feel that he is in a situation in which he will work with a professional person in a joint quest for understanding of himself, one can start with where the patient is. That is, the patient is asked why he thinks he is seeing a psychologist. Almost inevitably, there follows a discussion of what the problem is and the patient's feelings about it. The goal of the psychologist is to help the patient focus on the psychological aspects of his problem; after all, it is only for psychological problems that psychological testing is meaningful.

It is very frequent, especially in clinics and hospitals, for a patient to claim that he does not know why he is being seen. This generally represents a form of resistance and/or may be a part of rebellion against the person or persons instrumental in the referral, but it can also stem from poor preparation, such as a lack of information. Rather than ignore the purpose and meaning of this attitude by plunging in with an intellectual explanation of the nature of the contact, it is more important to use exploration in order to begin to focus the contact on the patient's psychological functioning. Thus, if the patient maintains that he does not know, one can raise the question of how come he complied with testing on such a basis, or did not seek out information, which in turn can open up the area of his passivity. This certainly occurs with children, who will claim that 'mother just brought them along and said they were to see a doctor, or perhaps that "my mother says I'm unhappy". The initial step is to get

at how they feel about this, does mother do this all the time, etc., aiming towards their feelings and difficulties in the parent-child relationship.

When the patient acknowledges a problem in the area of psychological functioning as the psychologist defines it, further exploration into what the patient thinks it is connected with is essential. One of the main reasons for this is that it may not be considered internal by the patient. This is not only important diagnostically, but the ability to view the problem as involving internal factors is also a necessary condition for the patient to see the value of testing for himself. While he may say that he always loses his temper, or that he cannot concentrate, he may attribute such difficulties to being picked on, or to a nagging wife or parent. In such cases, he is ultimately saying that he has no psychological problems, and that only the environment needs to be changed. If, on the other hand, there is some recognition of internal difficulties, testing for the purpose of clarifying his psychological functioning can be seen as obviously relevant.

Inter-related are the patient's feelings about his problems. Here we are in the area of motivation, which must be assessed. As in a pre-therapy interview, the nature of the presenting problem or symptom may be clarified, but it may also be obvious that there is no desire to change or that there is resistance to change. This attitude may then be explored, or may be set up as one of the problems to be worked on and understood; but, where the conditions allow, there may also be occasions when it appears best to help the patient make a decision about taking the test or not, and going along with his decision. Frequently, if he is very resistant, anticipates no benefit from testing, and perhaps even views testing as harmful, then even though he is compliant on the surface he will reveal little or nothing in

the tests. This occurs even when a patient is not deliberately withholding. In addition to the possibility of saving wasted time and avoiding more frustration for both testee and tester, it is necessary to recognize that there is no way to encourage positive feelings without being willing to elicit and face up to the negative ones, and to give people a choice.

The implication is not intended that every patient must agree that he has a problem, is unhappy about it, and wants help. Certainly, clinicians are generally acquainted with the patient who protests the loudest, and yet is most concerned: spending endless time trying to get him (in effect) to sign on the dotted line would be naive. Patients in places like corrective institutions, or mental hospitals, are not likely to show much self-awareness and are distrusting and resistant. While diagnostic work may still have to be undertaken in the face of these attitudes, it can be qualitatively different when the patients get some impression that their feelings are taken into account, and when the examiner faces realistically with the patient his negative attitudes towards testing and the necessity for proceeding despite them.

The most desirable attitude is one in which the patient shows some recognition of psychic upset and wants aid. With this attitude he will generally not only cooperate, but will even try to add as much as possible. Such an attitude is analogous to that of a person going voluntarily to a physician because of an illness; he not only gives his body over to the doctor, so to speak, but will frequently elucidate on his own, ask questions, mention possibly related self-observations. In short, he will at least attempt to be active and to participate with the physician, despite fears and resistances.

To elicit and maintain positive attitudes, it is also necessary to get at the patient's ideas concerning alleviation

of symptoms and/or complaints. Actually, the *main goal* if the patient has not come to this himself, is to communicate that the method for psychological help is predicated on understanding, and that this involves not only the understanding of the patient by the psychologist, but more importantly the understanding of the patient by himself in some degree or other. Only when a patient gets to this point, whether implicitly or overtly, can he realistically accept the diagnostic method, which, after all, is geared towards a fuller understanding. Moreover, as has been stated, an active participating attitude is encouraged when the patient grasps the value of his understanding, and goes beyond the attitude that he will simply serve as an object of study for the diagnostician.

Reaching this level of comprehension obviously sets the stage for introducing psychological tests. Ideally, we have a person who recognizes he has a problem or problems, wants help, and realizes that there is a need to know himself better. If a person is informed at all, he will generally have the notion already that tests contribute to understanding, that they "uncover", or at least verify or disprove; with motivation, such an individual will frequently seek out testing on his own, or welcome the opportunity. If the person is not informed, tests can be presented as a method for understanding him better and answering certain questions. It is advisable at this juncture to ask what the testee has heard of the various tests, in order to get at misconceptions. Everybody has some notions concerning tests, ranging from a child likening them to school exams to adults thinking that ink blots tell whether they are crazy or not. Simple, realistic explanations of the tests are best. Of course, these will only be accepted after feelings and distortions have already been handled. Thus, if a phobic child has associated seeing a psychologist with getting injections from a physician, no

amount of explanation that he is getting tests to understand his fears better will assuage his anxiety. On the other hand, if he is reassured that the psychologist does not use needles, then he will be more ready to listen and to assimilate an explanation of the tests and may even be eager to participate in getting to the bottom of his problem.

A patient who is encouraged to feel that he is an active part of the diagnostic process will rightfully expect some discussion of the results. Not only is this a general obligation to a person who has committed himself to being tested, but it also may have certain advantages. Certainly the psychologist cannot expect the patient to maintain an active cooperative attitude if this is to be one-sided only. The very expectation of the patient that the information he offers and the responses he gives will not be used "secretly" is often helpful in establishing a desirable attitude toward testing. Secondly, it is more often than not helpful to give a person at least some inkling of his problem areas. Too frequently we under-estimate patients in their ability to use such knowledge constructively. Also, further material may emerge in the course of interpreting findings, or the patient's reactions may serve to alter or even disprove certain initial results. Naturally, communication of findings has to be handled in such a way that it is at the patient's level of comprehension and that it is not too anxiety-provoking. One may not always be able to discuss findings with very young children, but rather with the parent, and here there must be assessment of how the results will be used.

In keeping with the underlying philosophy of this approach, the tests are considered as a part, rather than the whole, of the diagnostic process. Accordingly, when it is appropriate, one may use the test productions as they come up during testing to deal with resistances, focus on personality difficulties, and establish the condi-

tions under which a helping relationship is relevant. If during the Rorschach test, for example, it is seen that the patient is giving a guarded record, one may point out the patient's defensiveness and explore it. If a bright patient is failing at school because of emotional conflict, yet cannot recognize this, he may be confronted with his achievement on the intelligence test. The main point here is that the attempt to establish a meaningful test relationship is not confined to a "pre-test interview", but continues throughout the contact with the patient.

It should be made clear at this point that the testing procedures put forth in this paper are ideal goals or guides to be applied in all cases, including children of all ages, psychotics, delinquents, and mental defectives. This is in keeping with those moral and ethical principles which demand respect for the patient. Since this approach emphasizes the uniqueness and individuality of the patient, moreover, it must logically follow that one can advance no simple or rigid formulae which would apply to all cases. Nevertheless, it is felt that the principles outlined, used in accordance with sensitive clinical judgment, constitute a sound approach to the diagnostic goal, which is, after all, to understand and help the patient.

It may be of value, at this point, to list some of the things one does *not* do. One does not, for example, assume an attitude of indiscriminately cheerful friendliness to beguile the patient into cooperation. The psychologist does not artificially busy himself and thus avoid looking at and talking to the patient, on the grounds that the patient needs a moment or so to look him over. One does not trick the patient into cooperation; personality tests are not described as tests of "imagination". If the patient asks questions which are relevant to the testing, one does not put him off with promises to answer after testing is completed, hoping meanwhile that

he will forget his questions. One does not avoid facing the patient's anxiety; if the patient is worried about whether he is psychotic or not, and the question of psychosis is relevant, the psychologist does not tell him that testing has no bearing on the problem. One does not assume that there can be a single appropriate approach to all patients; there can be no cookbook prescription.

ILLUSTRATIVE CASES

Case I

The first case is illustrative of a rather common type of resistance, the manner of handling it, and the use of exploration and some interpretation during the tests.

C. was an adolescent girl referred for evaluation by the Juvenile Police because of a suicidal attempt after having engaged in sex play with a boyfriend. In the initial phase of his contact with her, C. was reticent about talking to the examiner. She had also been uncommunicative with the police. The psychologist had the impression that her distrust might be based on current experiences, and that it would result in her not responding to the tests. He reflected the girl's annoyance at having to repeat her story to still another person. He followed this up by explaining the clinic's and his functions, and their relationship to the Juvenile Police. She seemed somewhat relieved by this clarification. This was followed by asking C. what she thought of "doctors". She related past experiences of having been in the hospitals and having gotten needles and how frightened she was. The psychologist made a distinction between physicians and psychologists, and explained that his way of helping her was through understanding her better. She seemed to grasp this and poured out her story: her boy friend had attempted intercourse with her, she had refused, but later her friends accused her of having done it, called her names, and she became very despond-

ent. She could not understand how she had gotten into all of this, and her reaction was then used by the psychologist as the basis for testing, i.e. to understand why she had gotten so upset as to attempt suicide, and why she had difficulty in finding a better way to resolve her problems.

When testing was introduced, C. pointed to her head; this prompted the psychologist to explore the patient's concern that he was trying to find out if she were crazy. Noting her anxiety at this point, the exploration *was continued* with the resultant revelation by the patient that her mother had been mentally ill. A discussion of the girl's fears of her mother's behavior and her (C's) fear of the tests culminated in an agreement to take the tests, while being assured that she could discuss them with the psychologist.

Her Rorschach protocol was quite productive. A response involving a lonely animal prompted the examiner to ask C. about any feelings of loneliness she might have ever experienced. This query led the girl to reveal many such feelings revolving around her absent and ill mother, and in relation to this she revealed much affect by breaking into tears. At the same time the girl seemed to gain some insight; she commented that she had never quite thought of herself as lonely in quite that way before. At the conclusion of this appointment, C. indicated an eagerness to return, to take more tests, and to talk further with the psychologist.

Case II

This case illustrates the importance of using test productions when the interview does not seem effective in establishing an adequate relationship.

L. was a bright, alert looking nine year old who was brought to the Clinic because he had been in difficulty with the police as a result of some alleged fire-setting. Initially he presented the same type of picture

during testing that he presented when he was seen earlier for a psychiatric interview; that is, he spoke in positive terms of the many good things he had and did, and complained about almost nothing. He denied any feelings other than the general satisfaction with life. He denied setting the fire of which he was accused, but also denied any resentment about the accusation. He had spent about one month in a juvenile detention center, and L. said that this was "no place to send a boy for punishment" because it was too nice there. In response to additional questioning as to why he was at the Clinic, L. merely responded that he was coming under orders of the judge. Thus L., in his initial approach to the psychologist, appeared superficially to be quite intact, and the main reason for being referred was to assess whether serious pathology did, indeed, exist.

Following, and intermixed with this interview, psychological tests were administered to L. His protocols indicated that he was a severely disturbed boy, dangerously close to psychotic reaction; the test protocols contained a great deal of bizarre material. He seemed especially involved with fantasies about space men, monsters, bizarre anatomical notions, and the like. On the Rorschach he produced percepts with such content. An example is response to Card III of the Rorschach where L. produced the following percept: "Look like some people from Mars or space building another person. Building a strong man, or a girl or lady." It is necessary to reemphasize here that L. appeared clinically to be functioning in an intact manner and continued to deny any difficulty, worries or unhappiness about any aspect of his life. His ability to present a healthier clinical picture than the Rorschach was also reflected in L.'s showing more control and far less bizarreness on the TAT than he did on the Rorschach.

After the administration of the Ror-

schach, L.'s response to Card III was pointed out to him. The examiner inquired about whether L. often spent time thinking about such things as men from Mars. L. confirmed that he did and as the interview material developed, the earlier suspicion about L.'s engaging in isolated play and turning to fantasy living for satisfaction was confirmed. L. spoke about how much time he spent alone, either in movies or fishing or in parks. He also went on to elaborate other prominent fantasies which had not been reflected in the test material. One of these fantasies concerned his being a king, and there were others related to monsters. As L. became involved in relating these fantasies, the examiner suggested that in some ways it might be fun to think about these things, but that perhaps in other ways thinking about them caused L. a great deal of difficulty. For the first time L. showed what appeared to be a great deal of feeling and concern and responded, "yes, because sometime I get confused". The examiner inquired about this confusion and L. said that he imagined things and thought about them and then sometimes got mixed up about whether they were real or not. He then related a number of incidents where he imagined he saw a monster, and then got confused and did not know whether the monster was really there or not. On one occasion he was fishing when this happened, and literally threw his fishing pole into the air and ran from the park. He went on to say that he was now afraid to go into the particular park because of what had happened.

The examiner went on to inquire whether this being confused had something to do with L.'s getting into difficulty with the police, and L. said that it did. When asked in what way it had something to do with his getting into difficulty, however, L. did not respond. The examiner told him that he understood that this confusion could cause him a lot of trouble

and that it certainly caused a lot of worry, but that he did not have to talk about it at this time.

It was then suggested to L. that it might be helpful for him if he came to the Clinic and talked to someone about his confusion and these difficulties. L. was initially non-committal in response to this. The examiner went on to say that he understood how worried L. must be about this and how sometimes he is afraid and feels confused. He was told that sometimes coming to a place like the Clinic and talking with someone about things like this can be helpful and that we were interested in helping him. L. still did not respond to this. A few minutes later, however, on the way downstairs L. asked eagerly, when he was to come in again.

It should be noted that an initial attempt at exploring the patient's attitudes and feelings about the Clinic and his problems did not work, and the examiner tacitly accepted L.'s denial at the outset. However, the psychologist did not just accept the boy's productions but used them to penetrate the defensiveness. This in turn provided the psychologist with a wider sample of the boy's fantasy behavior and a more detailed look at the youngster's own reactions to his fantasies and anxieties. The assessment of the depth and extent of pathology was expedited, as was the patient's amenability to psychotherapy. Hence, the psychologist was in a relatively good position to answer the referral questions and in some detail. Further, the boy had experienced a situation closely paralleling therapy and it is likely that he was therefore better prepared to begin such a relationship. Lastly, it seems quite likely that the clarification and support he did experience during the testing itself might well have had some therapeutic benefits for him.

DISCUSSION

In the introductory section of this

paper, it was stated that the active handling of the test relationship has been found to have three broad consequences. Let us now re-examine these statements. Firstly, this type of interview-testing can be viewed as leading to a fundamentally more standardized situation. In terms of standardization by knowing and dealing with the patient's set, orientation and preconceptions, as pointed out for example, by Secord (1955) and Miller (1955), it is believed that the exploration done with the patient and the attempt to attain an "ideal relationship" do contribute to a more uniform diagnostic situation from patient to patient.

A second advantage is that of potentially making available more meaningful material, and thus increasing clinical validity. In their book on Rorschach interpretation, Phillips and Smith (1953) point out that guarded, sparse protocols do not contain "differentiating personalized material" but only "the popular and conventional" (p.181). "It is from unconventional responses and from elaborations of both conventional and unconventional responses that individuating characteristics are inferred. And it is these . . . which are absent from the guarded records" (p.183). Thus, these authors explicitly recognized the diagnostic problems posed by such test data, but they do not go beyond this to suggest how the examiner might elicit a more useful record. It is felt that perhaps the most significant advantage occurring from the active handling of the patient-psychologist relationship is increased productivity, and thus a broader base from which to approach the diagnostic problem.

Not only may more detailed protocols be obtained, but the psychologist has the opportunity to observe firsthand the patient's usual modes of interpersonal interaction and defense. Further, the additional material gained from the patient's own ideas about his productions may give more

cues upon which to base interpretive hypotheses. Indeed, these cues would also seem to aid in narrowing down the number of interpretations to be considered for any given set of examination data. Not less important is the opportunity the clinician has to assess the patient's strengths, his ability to adopt new ways of thinking about himself, and his amenability to treatment.

Finally, such a patient-(and feeling)-oriented testing experience helps to introduce the patient to clinical help. In this cooperative work, the patient has a demonstration of how one looks at oneself, how one begins to notice the unnoticed, remember the forgotten, relate the unrelated. In short, he is provided with an experience which is in many respects a paradigm of the therapeutic hour. The transition from appraisal period (i.e., diagnostic work-up) to treatment is then smooth and natural. This last factor is rarely accorded the importance it deserves. Not infrequently the psychological testing may be the patient's first or only contact with the clinic; it may be critical in determining motivation and attitudes toward further contacts, or perhaps the insight and clarification gained during this contact may be all that the patient will receive. It behooves us, then, to make the experience as positive and beneficial for the patient as possible.

The position is not taken that this technique is without disadvantages or that it is fool-proof. It is clearly recognized that it works with varying success with different patients, that it requires a good deal of skill by the psychologist, that it embodies potential dangers to some patients, and that it is very difficult to teach. Further, one may not wish to emphasize the therapeutic aspects equally in all clinical settings.

Despite these shortcomings, however, the fundamental advantages of psychological testing conducted in the manner proposed seem so strongly

supported by experience as to warrant continued study and refinement of this approach.

REFERENCES

- Bell, J. E. *Projective techniques*. New York: Longmans, Green & Co., 1948.
- Calden, G., & Cohen, L. B. The relationship of ego-involvement and test definition to Rorschach test performance. *J. proj. Tech.*, 1953, 17, 300-311.
- Cronbach, L. J. *Essentials of psychological testing*. New York: Harper & Bros., 1949.
- Gibby, R. G., Miller, D. R., & Walker, E. L. The examiner's influence on the Rorschach protocol. *J. consult. Psychol.*, 1953, 17, 425-428.
- Holt, R. R. Gauging primary and secondary processes in Rorschach responses. *J. proj. Tech.*, 1956, 20, 14-25.
- Hutt, M. L. A clinical study of "consecutive" and "adaptive" testing with the Revised Stanford-Binet. *J. consult. Psychol.*, 1947, 11, 93-103.
- Hutt, M. L., Gibby, R. G., Milton, E. O., & Pottharst, K. The effect of varied experimental "sets" upon Rorschach test performance. *J. proj. Tech.*, 1950, 14, 181-187.
- Joel, W. The interpersonal equation in projective methods. *Rorschach Res. Exch.*, 1949, 13, 479-482.
- Klatskin, Ethelyn H. An analysis of the effect of the test situation upon the Rorschach record: formal scoring characteristics. *J. proj. Tech.*, 1952, 16, 193-199.
- Klopfer, B. & Kelley, D. M. *The Rorschach technique*. Yonkers-on-Hudson: World Book, 1946.
- Lord, E. Experimentally induced variations in Rorschach performance. *Psychol. Monogr.*, 1950, 64, 1-33.
- Luborsky, L. Self-interpretation of the TAT as a clinical tool. *J. proj. Tech.*, 1953, 17, 217-223.
- Luft, J. Interaction and projection. *J. proj. Tech.*, 1953, 17, 489-492.
- Miller, D. R. Prediction of behavior by means of the Rorschach test. *J. abnorm. soc. Psychol.*, 1953, 48, 367-375.
- Phillips, L. & Smith, J. G. *Rorschach interpretation: advanced technique*. New York: Grune & Stratton, 1953.
- Sarason, S. B. *The clinical interaction*. New York: Harper & Bros., 1954.
- Schachtel, E. G. Subjective definitions of the Rorschach test situation and their effect on test performance. *Psychiat.*, 1945, 8, 419-448.
- Schafer, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- Schafer, R. Transference in the patient's reaction to the tester. *J. proj. Tech.*, 1956, 20, 26-32.

Secord, P. F. An analysis of perceptual and related processes occurring in projective testing. *J. gen. Psychol.*, 1953, 49, 65-85.

Sinick, D. Anxiety in the testing situation. *Personnel Guid. J.*, 1953, 31, 384-387.

Terman, L. M., & Merrill, Maud A. *Measuring intelligence*. Boston: Houghton Mifflin, 1937.

Wells, F. L. *Mental Tests in clinical prac-*

tice. Yonkers-on-Hudson: World Book, 1927.

Wechsler, D. *The measurement of adult intelligence*. Baltimore: Williams & Wilkins, 1939.

Wolberg, L. R. *The technique of psychotherapy*. New York: Grune & Stratton, 1954.

Received May 3, 1961

Interlevel Disparity and Predictive Efficiency¹

JAMES C. MCGREEVEY

Clark County Mental Health Center, Vancouver, Wash.

There has been an increasing awareness of the problem of "levels" of personality throughout the history of personality theory, with a gradual shifting of emphasis from interest in the individual, or non-social aspects of personality, to interest in man involved in interpersonal relations (Benedict, 1934; Mead, 1935; Kardiner, 1945; Linton, 1945; Horney, 1945; Fromm, 1947; Sullivan, 1947).

Freedman and Leary (1951) discuss the need for a comprehensive scheme for description of personality, and the necessity that the variables being studied possess interpersonal reference. The interpersonal approach to personality has also been expanded by Leary in his book (1957). Leary defines personality as the multilevel pattern of interpersonal responses expressed by the individual, and suggests that all social, emotional and interpersonal activities of an individual can be understood as attempts to avoid anxiety or to establish and maintain self-esteem. Leary suggests five levels of personality. The present study is concerned with the first three levels: I. The Level of Public Communication, which consists of ratings by others of the overt behavior of an individual; II. The Level of Conscious Description which includes the person's description of himself and others, or the way he reports his views of himself and the world; III. The Level of Private Symbolization, which is the indirect expression of an individual's imagined self and relationship in his preconscious or symbolic world.

It is evident from disagreements in the literature regarding the accuracy with which projective tests predict behavior and the "levels" of behavior being predicted that there is a need for a closer equating of inferences from tests with specified "levels" of personality.

Many psychologists would agree with the statement of Murray (1951): "Whatever particular virtue the TAT may have, if any, it will be found to reside, not as some have assumed in its power to mirror overt behavior, but rather in its capacity to reveal things that the patient is unwilling to tell or is unable to tell because he is unconscious of them." Yet Allport (1953) in his critique of projective testing suggests: "... normal subjects ... tell you by the direct method precisely what they tell you by the projective method ... you may therefore take their motivational statements at their face value, for even if you probe you will not find anything substantially different."

PROBLEM

The present study is concerned with the need for equating inferences from tests with "levels" of personality by investigating the use of self-rating questionnaires and projective tests for predicting Level I behavior (the way a person is perceived by others) and Level II behavior (the way the person perceives himself), and with the effect exerted on such predictions by the presence or absence of ego-threat. It is assumed that a certain amount of defensiveness is present for the purpose of maintaining self-esteem when a person directly ranks himself on a trait and that the amount of defensiveness varies with the presence or absence of ego-threat. That is, the defensive person will tend to over-rate

¹ This study is based upon a dissertation submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy at the University of Portland under the direction of a committee composed of Professors Walter G. Klopfer, Gordon K. Higginson, William A. Botzum, C.S.C., and Sheridan P. McCabe.

himself on desirable traits he does not possess according to others, and to under-rate himself on undesirable traits he is considered to possess. This defensiveness should operate to a greater extent when a person reports his self-concept by direct comparison with others than when he answers questions on a self-rating questionnaire type test or when he responds to projective tests. This would appear to be plausible because, in the instance of questionnaire and projective tests, the person is less aware of the dimension being investigated and thus is less able to be defensive. When ego-threat is present, predictions from projective tests should correspond more closely to Level I behavior than to Level II behavior, since perceiving another as possessing undesirable traits is less ego-threatening than perceiving one's self as possessing these traits. With non-ego-threatened subjects, however, predictions from projective tests should correspond similarly to both Level I and Level II behavior.

In the present study, inter-level disparity or the degree of disparity between Level I and Level II behavior is considered to be a measure of ego-threat. Inter-level disparity here refers only to differences resulting from the subject's being ranked by peers as possessing undesirable traits or as lacking desirable traits, and the subject ranking himself oppositely. Ego-threat is thus defined operationally as being present when he ranks himself high on desirable traits which others feel he lacks or low on undesirable traits which others feel he possesses. The study proposes to test the following four experimental hypotheses:

1. Self concept or Level II measures of personality traits will not differ significantly from questionnaire test measurements of the same traits for a non-ego-threatened group but will differ significantly for an ego-threatened group.

With the ego-threatened group, it is predicted that defensiveness will

result in the subjects ranking themselves higher on desirable traits and lower on undesirable traits when reporting their self concept. Since the direction of differences is being predicted for the ego-threatened group, a one tailed test of significance will be used in testing the second part of this hypothesis.

2. Peer concept on Level I measures of personality traits will not differ significantly from questionnaire test measurements of the same traits for either an ego-threatened or non-ego-threatened group.

3. Self concept on Level II measures of personality traits will not differ significantly from clinician's judgments of the same traits from projective tests for a non-ego-threatened group, but will differ significantly for an ego-threatened group.

As in the case of hypothesis I above the direction of difference is being predicted for the ego-threatened group and a one-tailed test of significance will be used in testing the second part of this hypothesis.

4. Peer concept of Level I measures of personality traits will not differ significantly from clinician's judgments of the same traits from projective tests for either an ego-threatened or non-ego-threatened group.

METHOD

Third and fourth year student nurses were used as subjects. Because the subjects were required to rank each other and ranking a large group is difficult, the data were obtained from three separate groups numbered 22, 23, and 26 Ss.

Self-ranks, ranks by peers, objective test scores and predictions by clinicians from projective tests were obtained for the 71 Ss on two desirable and two undesirable traits. The two desirable and two undesirable traits were selected from the ten categories of the Mental Health Analysis (MHA) (Thorpe and Clark, 1946). The selection was based on rankings made by third and fourth year college

students on the desirability of the traits. The two ranked most desirable and the two ranked least desirable were selected. The desirable traits are Adequate Outlook and Goals (AOG) and Close Personal Relationships (CPR). The two undesirable are Emotional Instability (EI) and Behavioral Immaturity (BI).

From the combined subjects, two groups of 20 Ss each were selected separately for each of the four traits, one group having the highest disparity between self rankings and peer rankings; the second having the lowest disparity. The high disparity group is operationally defined as the ego threatened group, and the low disparity group as the non-ego-threatened group.

The MHA was administered for each of the four traits to obtain a self-ranking questionnaire test measurement. Raw scores from the MHA were converted to ranks which were used in the analysis of the data.

Projective tests, consisting of the Sacks Sentence Completion Test (SCT) and five selected Thematic Apperception Test (TAT) cards were then administered.² Each S was then asked to rank all other Ss in her group on each of the four traits from the MHA. The Ss were finally asked to designate the amount of the trait they felt they possessed by indicating their own position in the ranks.

The SCT and TAT were administered in group form. Lindzey and Heinemann (1955), in comparing group and individual administration of the TAT, report no consistent pattern favoring either method. When the two methods were compared in

terms of relationship to independent measures (questionnaires, self-ratings, sentence completion test) there was slight evidence suggesting superiority of the group method.

To obtain clinicians' ranks on each trait six clinical psychologists possessing a Ph.D. degree and with at least five years experience in the use of projective tests were used as judges.³ Three of the clinicians ranked the Ss from the TAT protocols and the other three ranked from the SCT. A counterbalancing order was used so that each clinician ranked two of the original three groups of Ss on all four traits. Thus, four clinicians ranked each group, two using the TAT protocols, the other two using the SCT. The data were analyzed for the pooled judgments from the TAT and SCT together. The ranks assigned to each group were the mean of the four clinicians who ranked the group.

The use of pooled judgments from TAT and SCT together is supported by Little and Shneidman (1959). They state that agreement among test interpreters is not to be expected when tests are used singly. However, they point out the possibility that combinations of projective tests might produce better agreement. The use of pooled judgments is further supported by Silverman's (1959) Q-Sort study of the validity of evaluations made from projective techniques. Silverman found that clinicians were able to appraise projective tests with a degree of agreement significantly greater than chance when their appraisals were based on their total impressions from a number of tests.

RESULTS

To determine the extent of agreement among clinicians' judgments, rank order correlations were computed between the mean rank of the two clinicians who judged each group

² TAT Cards 1, 2, 3BM, 3GF, and 6BM were used. These cards were selected by having two clinicians rank 12 cards separately for each of the four traits of the MHA according to the best card for depicting the trait in question. It was originally planned to select the two cards ranked best for each trait or a total of 8 cards. However, because three cards were ranked best for both of two traits, the number of cards was limited to five.

³ The author is indebted to Sidney F. Dean, Gordon K. Higginson, Bernard I. Murstein, Allen H. Parker, Sue A. Warren, and Arthur W. Wiens who contributed their time and skill in completing the clinicians' judgments.

from the SCT and the mean rank of the two who judged from the TAT. These results are presented in Table I.

Although the correlations in Table I are not as high as would be desired, the majority are significant beyond the .01 level. The three correlations which do not reach the .05 level do approach significance in the desired direction.

Table II presents a comparison of MHA scores with self ranks and with peer ranks for both the low and high disparity groups.

According to hypotheses 1 and 2 it was predicted that the self ranks for the high disparity group would differ significantly from the MHA scores, but the peer ranks for the high dis-

parity group and the self and peer ranks for the low disparity group would not differ significantly. The results supported the first two hypotheses with the self ranks and questionnaire test differing significantly on three of the four traits at the .05 and .01 level for the high disparity group, while a significant difference occurs on only one of the remaining 12 comparisons. Inspection of the data shows that the differences between the MHA scores and the self ranks for the high disparity group are in the direction of greater possession of desirable traits and less possession of undesirable traits for the self concept than the test scores.

Table III presents a comparison of

TABLE I—Correlations between Means of Clinicians' Ranks from Thematic Apperception Test and Sentence Completion Test for Group A, B and C Subjects

MHA Traits	Group A	Group B	Group C
	Av. of Clinicians 182 vs. 485	Av. of Clinicians 183 vs. 486	Av. of Clinicians 283 vs. 586
AOG	.56**	.36	.24
CPR	.71**	.57**	.37
BI	.64**	.54**	.49*
EI	.69**	.66**	.51*

* .05 level of significance

** .01 level of significance

TABLE II—Rank Test for Comparison of Mental Health Analysis with Self Ranks and with Peer Ranks for Low and High Disparity Groups

MHA Traits	Low Disparity Group		High Disparity Group	
	Self	Peer	Self	Peer
AOG	1.12	1.53	2.35**	1.75
CPR	.09	.52	1.16	.97
BI	.78	1.42	2.03*	1.34
EI	2.22*	1.31	2.15*	.84

* one tail test of significance

* .05 level of significance

** .01 level of significance

TABLE III—Rank Test for Comparison of Self and Peer Ranks with Pooled Clinicians' Ranks for Low and High Disparity Groups

MHA Traits	Low Disparity Group Clinicians With		High Disparity Group Clinicians With	
	Self	Peer	Self *	Peer
AOG	1.64	.93	3.92***	1.15
CPR	.63	1.53	3.92***	.35
BI	1.10	.48	2.50**	2.13*
EI	.69	.45	3.21***	.52

* one tail test of significance

* .05 level of significance

** .01 level of significance

*** .001 level of significance

clinicians' pooled ranks from the TAT and SCT with self and peer ranks for the low and high disparity groups.

As predicted in hypotheses 2 and 3 the self ranks and clinicians' judgments differ significantly on all four traits for the high disparity group. Neither the self nor peer ranks for the low disparity groups are significantly different from the clinicians' judgments and for only one of the four traits are the peer ranks significantly different for the high disparity group. The differences between self ranks and clinicians' judgments for the high disparity group are in the direction of self ranks reporting greater possession of desirable traits and less possession of undesirable traits.

DISCUSSION AND CONCLUSIONS

A relatively high agreement was obtained between the clinicians' judgments from the TAT and SCT of the four traits involved in the present study. Significant correlations were obtained on 9 of the 12 comparisons made. However, a greater extent of agreement than obtained would be desirable, and might be expected of clinicians when making judgments regarding personality traits from projective tests. The difficulty does not appear to lie in the area of degree of training or skill on the part of the clinicians since care was taken to select highly trained judges with considerable experience in the use of projective tests. It would seem, therefore, that the weakness lies in the lack of objective criteria for making judgments of this nature. A need for further study in objectifying the criteria from which judgments are to be made from projective tests is indicated from these results.

The results obtained in this study indicate that questionnaire test scores and self ranks do not differ significantly when the person is non-ego-threatened with regard to the trait being measured, but do differ significantly if ego-threat is present. The defensiveness of the ego-threatened

person appears to be more manifest on direct reporting of the self concept than when answering items on the questionnaire test. The self concept is reported as including a greater amount of the desirable traits and a lesser amount of the undesirable traits. Thus, the person reports himself to others in such a manner as to enhance self-esteem. On the other hand, peer concepts did not differ to a significant degree from the questionnaire test scores for either the ego-threatened or non-ego-threatened group. It would appear then that questionnaire tests predict the way the person is seen by others as well for an ego-threatened group as for a non-ego-threatened group, but do not predict the way the person sees himself as well for one group as for the other.

Comparison of self ranks with the pooled ranks by clinicians supports the third hypothesis by showing that the self ranks for the non-ego-threatened group are not significantly higher or lower than the clinicians' ranks, and that the self ranks for the ego-threatened group are significantly different from the clinicians' ranks. This difference is in the direction of self ranks signifying greater possession of desirable traits and less possession of undesirable traits. This indicates that when clinicians predict personality traits from projective tests for a non-ego-threatened group, their predictions are not consistently higher nor lower than the self-concept. However, for an ego-threatened group, their predictions differ from the self concept in a consistent direction and to a significant degree. These findings would tend to support the claim of Murray (1951) that the value of the TAT resides in its capacity to reveal things that the patient is unwilling to tell, and to disagree in part with Allport's (1953) contention that normal Ss tell you by the direct method precisely what they tell you by the projective method. In view of the present findings, it seems that Allport's statements might be modified as fol-

lows: Non-defensive Ss tell you by the direct method precisely what they tell you by the projective method, but defensive Ss tell you something significantly different by the direct method than they do by the projective method. Such a modification of Allport's statement assumes that defensiveness and normality are not mutually exclusive, and that normal Ss may or may not be defensive in regard to any given trait.

Comparison of peer ranks with the pooled ranks by clinicians supports the fourth hypothesis. Peer concept or Level I measures of personality traits did not differ significantly from clinicians' judgments on any of the four traits for the non-ego-threatened group and on only one of the four traits for the ego-threatened group.

The over-all results of the study suggest that ego-threat is a phenomenon found among clinically normal Ss and can be independently measured for various traits. It is defined as the denial of undesirable traits perceived by others and claimed desirable traits not perceived by others. This ego-threat is manifested as defensiveness which influences the self concept in the direction of greater self esteem. In the case of ego-threatened traits, self ranks cannot be predicted accurately from questionnaire test scores or from clinicians' judgments of traits from projective tests. When ego-threat is absent or when peer judgments are used, this defensiveness is not operative, and neither the self concept nor the peer judgments differ significantly from either questionnaire test scores or clinicians' judgments from projective tests.

The results of the present study have implications regarding the level of behavior which can be more accurately predicted from projective tests. It would appear that Level II behavior for a high disparity or ego-threatened group cannot be predicted as accurately as Level II behavior for a non-ego-threatened group, and that Level II behavior for an ego-threat-

ened group cannot be predicted as accurately as Level I behavior for either an ego-threatened or non-ego-threatened group. That is, clinicians can most safely predict social behavior (the way the person is seen by others), but can safely predict conscious behavior (the way the person sees himself) only for non-ego-threatened areas. This implies that some of the negative or contradictory results in studies with projective tests using Level II behavior as criteria may in part be due to some of the traits measured being high disparity traits and some being low disparity traits, which could operate to lessen the likelihood of achieving reliable predictions.

The results of this study also appear to have bearing on validity studies with projective tests. Since the presence or absence of ego-threat appears to affect predictions of Level II behavior, but does not appear to affect the predictions of Level I behavior, the implication is that more emphasis might be placed in using Level I behavior as a criterion in validity studies.

Leary (1957), in describing Level III or "preconscious" behavior states:

"In actuality subjects do not always express in response to projective stimuli their private or fantasy thoughts. Many defensive, suppressive, rigid patients repeat in their response to projective stimuli the same themes they report in their conscious description. What we get at Level III is therefore not always preconscious material, but those themes which the subject is willing to express in the testing situation."

There is an apparent contrast between the above statement of Leary's and the findings of the present study. In the present study, defensiveness has been defined as Level I-II discrepancy, and it has been found that this condition is more inclined to be associated with Level II-III discrepancy rather than Level II-III congruence as implied by Leary.

The present results suggest that Clinicians' predictions from projective tests or Level III measures of behavior

are congruent with the subjects' conscious descriptions for only non-defensive areas, i.e., those where there is little disparity between Level I and Level II behavior. However, this situation does not appear to exist for subjects who are defensive in regard to a given trait, i.e., for those with large disparity between Level I and Level II behavior. For these Ss, clinicians' predictions from projective tests appear to be more congruent with Level I behavior or the way the person is seen by others.

The difference between clinicians' predictions from projective tests and the Ss' conscious descriptions which was found to exist for the defensive Ss might be due to a greater freedom for defensiveness to operate in the case of self-ranking than when the S responds to projective material. When the S responds to projective material he is not as aware of the meaning of his responses as they will be interpreted by the clinician as he is when he reports directly his self concepts. The greater subtlety of the questionnaire approach results in a similar diminution of the effects of defensiveness. For Level I measures of behavior and for self-ranks with only non-defensive Ss this condition does not appear to exist because of the absence of need to maintain self-esteem.

It can be seen from the above that Level II behavior appears to be congruent with Level III behavior only when there is little disparity between Level I and Level II. Level I behavior does not appear to differ significantly from Level III behavior regardless of the difference between Level I and II. It would appear that further investigation of Level II behavior and of the effect of differences between Level I and Level II behavior might be helpful in achieving a clearer understanding of Level III or projective measures of behavior, particularly as the latter is utilized by clinicians in the prediction of various levels of behavior.

SUMMARY

The present study was concerned with investigating the accuracy of self-rating questionnaire tests and projective tests for predicting Level I and Level II behavior as defined by Leary, and with the effect expected on these predictions by the presence or absence of ego-threat. Self concept and peer concept were compared with the questionnaire test scores and with clinicians' judgments from projective tests for both ego-threatened and non-ego-threatened groups. From the results of these comparisons the following conclusions are suggested:

1. Self concept or Level I behavior can be predicted more accurately from questionnaire tests for non-ego-threatened traits than for ego-threatened traits.
2. Peer concept or Level II behavior can be predicted as accurately from questionnaire tests for ego-threatened traits as for non-ego-threatened traits.
3. Clinicians' judgments from projective tests agree more closely with self-concept for non-ego-threatened traits than for ego-threatened traits. With ego-threatened traits, clinicians' judgments disagree with self concept to a statistically significant degree and in a consistent direction, i.e., under-rating desirable traits and overrating undesirable traits.
4. Clinicians' judgment from projective tests do not agree more closely with peer's concept for non-ego-threatened traits than for ego-threatened traits.

The results of the study were discussed with reference to the levels of behavior which can most accurately be predicted from projective tests, the contradictory results in studies predicting Level II behavior from projective tests and the implication of the results for validity studies with projective tests.

REFERENCES

- Allport, G. W. The trends in motivations theory. *Amer. J. Orthopsychiat.*, 1953, 23, 108-119.

- Benedict, R. *Patterns of culture*. Boston: Houghton Mifflin, 1934.
- Freedman, M. B. & Leary, T. F. The interpersonal dimension of personality. *J. Pers.*, 1951, 20, 143-161.
- Fromm, E. *Escape from freedom*. New York: Rinehart, 1947.
- Horney, Karen. *Our inner conflicts*. New York: Norton, 1945.
- Kardiner, A. *Psychological frontiers of society*. New York: Columbia University Press, 1945.
- Leary, T. *Interpersonal diagnosis of personality*. New York: Ronald Press, 1957.
- Lindzey, G. & Heinemann, S. H. Thematic Apperception Test: Individual and group administration. *J. Pers.*, 1955, 24, 35-55.
- Linton, R. (Ed.) *Cultural background of personality*. New York: Appleton-Century-Crofts, 1945.
- Little, K. B. & Shneidman, E. S. Congruencies among interpretations of psychological tests and manometric data. *Psych. Monographs*, 1959, Whole No. 6, 73, 42.
- Mead, M. *Sex and temperament in three primitive societies*. New York: William Morrow, 1935.
- Murray, H. A. Uses of the Thematic Apperception Test. *Amer. J. Psychiat.*, 1951, 107, 377-381.
- Silverman, L. H. A Q-Sort study of the validity of evaluations made from projective techniques. *Psych. Monographs*, 1960, 74, Whole No. 7, 28.
- Sullivan, H. S. *Conceptions of modern society*. Washington, D. C.: William Alanson White Psychiatric Foundation, 1947.
- Thorpe, L. P. & Clark, W. W. *Manual of directions. Mental Health Analysis, Adult Series*. California Test Bureau, 1946.

Received August 1, 1961

An Indirect Validation Study of the Draw-A-Person Test Through the Cartoons of William Steig

DONALD MCGUIRL
State Hospital No. 1, Fulton, Mo.

C. SCOTT MOSS
National Institute of Mental Health

INTRODUCTION

In a recent article, Sundberg (1961) presented data collected from a survey of 185 hospitals and clinics indicating that after the Rorschach, the Draw-A-Person Test (DAP) is the most frequently used psychological test in the United States today; the author also distinguished between highly researched and under-researched tests, citing the two as respective examples. The DAP has achieved this level of preeminence despite the fact that few research evaluations have supported its reliability and validity. Swensen (1957) in a review of the literature through 1956, found that only one interpretative hypothesis had suggestive support. A survey of the literature through 1960 (Goldschmid) revealed 21 additional studies, the majority of which were again not encouraging. However, two points need to be emphasized: (a) few experimental evaluations have been conducted on this test relative to the Rorschach and many other projective methods, and (b) the reliability and validity of the DAP is not demonstrably lower than other projective techniques, most of which have proven resistant to traditional methods of assessment.

Machover (1955) in an article dealing with the role of body image in art communication, singled out the cartoonist, William Steig, as an individual possessed of a singular ability to communicate a wide range of feeling, states of mind, and characterological defenses through an inspired sensitivity for the "language the body speaks." In this article, she examined a selection of 36 Steig drawings which she felt could "almost serve as illustrations for clinical elucidation of problems of interpretation of the drawing

of the human figure in particular, and the psychology of expressive movement in general." Machover saw the cartoons as very rich in cues important for the fine differentiation of behavior traits and expressiveness of others.¹

(Steig) digs deep for the causes, multiple manifestations, and the total complexity of body experience . . . Through the plastic manipulation of the human form, at which Steig is an expert, he manages to depict graphically the most complicated factors in the anatomy of human emotions . . . Steig conveys bigness, appropriate placement of the figure on a page, the pressured line of motor intensity, the more torturous and lighter line of the obsessive thinker, the concealment and facade values of clothing, and universally experienced sexual symbols in ways not unsimilar to the projections of any individual who attempts to "draw a person" (pp. 456-457).

In this article Machover utilized many of the hypotheses advanced in her *Personality Projection in the Drawing of the Human Figure* (1949) to give a detailed clinical interpretation to these drawings.

Steig's cartoons would thus seem to provide stimulus material particularly rich in many of the qualities defined by Machover as contributing to accuracy in judging others through her test. Machover further contends:

Steig articulates graphically many minute and intimate body experiences . . . Identification with the drawings, although at times uncomfortable, is spontaneous, although we do not fully know why. The fact remains that the detailed tracing of

¹Some of the drawings were three dimensional and shaded, with the body forms realistically clothed; others were line drawings emphasizing, in symbolic fashion, various feelings and attitudes; and one drawing was highly abstract, though Machover regarded its meaning readily communicable.

our common body experiences, consciously executed by Steig, strikes a chord of intimate familiarity . . . (He) is an extremely serious artist who talks with *his* body, about *our* bodies . . . The viewer of Mr Steig's drawings experiences a "just right" feeling, although hardly understanding the nature of the intimacy. It is as if he, the viewer, would draw that way, if he felt the way that Steig's people are portrayed (pp. 455-460).

From such assertions, it would be concluded that almost everyone understands the "body-language" utilized by Steig, although perhaps not knowing why—that Steig's singular ability to communicate through this media is in fact, the basis for his popularity as a cartoonist. This understanding then should be demonstrated to a significant degree throughout the population, regardless of differences in age, experience, training, intelligence, and other factors. It was the purpose of the present study to test this hypothesis.

METHOD

As the basic test group, it was felt that persons should be chosen who would represent the "man-in-the-street," that is, individuals with an average amount of intelligence, education, and normal use of their faculties. If this group demonstrated an understanding of Steig's drawings, the hypothesis would be regarded as tenable. As an extension of the study, two "expert" groups were also selected who were assumed to represent atypical degrees of experience or training in the utilization of non-verbal cues. The first consisted of congenitally deaf individuals, and the second a population of clinical psychologists. The performance of these groups was expected to reflect the effects of: (a) special dependence on non-verbal cues for an understanding of the environment; and, (b) specialized training in the evaluation of personality and in the use of the DAP. In addition, the contributions of general intelligence and confidence were examined.

Selection of Subjects

Specifically, the three main groups of subjects in this study were as follows: A Normal Group, recruited from a non-college population; a Deaf Group, selected from a population of young adults in a school for the deaf; and a group of practicing Ph.D. clinical psychologists. The 20 normal subjects had a mean age of 32 and a wide range of occupations were represented. Two subgroups were defined on the basis of intelligence using a short form of the Wechsler Adult Intelligence Scale. Subgroup A was composed of 10 persons whose estimated I.Q.'s fell within the range of 90-110; those in Subgroup B had I.Q.'s of 120 or above.²

There were also 20 subjects in the deaf group drawn from a population of high school juniors and seniors between the ages of 16 and 20. Originally, it had been intended that only those persons would be selected whose Full Scale WAIS I.Q.'s fell between 90 and 110, or above 120. However, none of the individuals from this population attained Full Scale I.Q.'s in the Superior range.³ As far as practicable, individuals were selected who had suffered their hearing loss before

²The short-form consisted of the two sub-scales of Information and Similarities, which have been shown to correlate .903 with the full scale when administered as a diad (Maxwell, 1957).

³It is well known that one of the chief impairments among the hearing deprived is the retarded development of language and verbal facility, an impairment almost always reflected on the WAIS by a superiority of the Performance Scale over the Verbal Scale I.Q. The largest single discrepancy found in the present group was a PSI.Q. of 126 as contrasted with a VSI.Q. of 70. In order to facilitate comparison between the Deaf and Normal Groups, a prediction in the former group of the Full Scale I.Q. was made from the Information and Similarities subtests alone and the group was divided on this basis. Thus the Deaf and Normal Groups were not matched on intelligence; however, there was a within group relativity. It may be objected that prediction from these admittedly depressed subtest scores is questionable, but the predicted I.Q. correlated .808 with the actual FSI.Q.

the age at which substantial language facility could develop. Seven of the subjects were born deaf, and only two were over the age of one and a half years when deafness occurred. The present degree of hearing loss ranged from "marked" (over 40 decibels), to "profound" (80 decibels and over). In practical terms, only three of these individuals retained sufficient hearing to understand conversation even when shouted.

Ten clinical psychologists ranging in age from 29 to 52 years constituted the third group. All had from two to 12 years of post-doctoral experience. They had familiarity with the DAP Test, and most had utilized this instrument frequently in their work. No attempt was made to control for intelligence in this group; it seemed reasonable to assume that they would all belong to the superior intelligence category.

Materials

The test consisted of reproductions of 32 Steig cartoons, and 32 tabs consisting of the printed captions originally appended to the cartoons by the artist. It was not possible to utilize all 36 of the cartoons cited by Machover in her article since the artist left four of these uncaptioned.⁴ In the photographic process, the proportions of all the drawings were equalized so they could be centered on $4\frac{1}{2} \times 3\frac{1}{2}$ " cards.

Wherever possible the captions were worded in precisely the original manner. One exception occurred with the caption "Weltschmerz," a German noun which is defined as, "pessimistic outlook on life in general." In this case both the German title and the English equivalent were typed on the tab. Most of the captions were simi-

larly descriptive of feelings, states of mind, or character defenses. In addition, some of the words (e.g., "serene," "gregarious," "affable") were considered beyond the vocabulary level of some of the deaf and normal hearing group. In an attempt to equalize as much as possible the verbal stimuli, definitions of the more difficult words were typed below the original captions. As a precaution against confusion among the deaf subjects, some of the original captions were also paraphrased into more concrete terminology (e.g., original: "pessimistic outlook" — paraphrase: "thinking unhappy about the future"). When used the paraphrases were typed at the bottom of the tab.

Procedure

The experimental procedure involved asking the subjects to study the Steig drawings and to match them with a verbal expression (caption) which they felt best defined its meaning. All cartoons and captions were arrayed on a table in front of the subject and he was free to alter his matchings until he felt satisfied. There was no time limit. When he had completed the task, each subject was asked to select the 10 matchings he felt most confident were correct, and the 10 of which he felt least confident.

A slight modification was necessary in the case of the deaf subjects. To facilitate their understanding of the task, an interpreter, experienced in sign language, was present to explain to them what they were to do.⁵ Furthermore, the deaf subjects were allowed to ask for a sign language translation of a caption word or phrase at any time if they were in doubt as to its meaning.

RESULTS

From a total of 1600 matchings, 263

⁴All 36 Steig cartoons are illustrated on pages 454 and 455 of the Machover article. Cartoons designated Figures XIV, XV, XXI, and XXIV were excluded. In a personal communication, the artist indicated that he preferred to let these particular cartoons "speak for themselves."

⁵For serving as interpreters in this phase of the study the authors are indebted to Dr. C. P. Goetzinger and Mr. Dennis Ortiz. Appreciation is also expressed to Arthur Roehlke for his assistance in the original formulation of this design.

(16 per cent) were in agreement with the meanings intended by Steig. The number of correct matchings per individual ranged from zero to 12, with an average of 5.25. Table I shows the significance of the performance of all 50 subjects, and for each of the experimental groups. As is indicated, the difference between the expected and the observed number of correct matchings is highly significant ($p = < .0001$). Thus, the basic hypothesis that meanings can be communicated effectively through Steig's drawings was firmly supported.

TABLE I. Significance of the Observed Number of Correct Matchings for all Subjects and for Each Experimental Group (Chance Probability is either Zero or One Correct Choice)

Group (N)	Number Exceeding Chance	Probability
Normal-A (10)	9	<.0000413
Normal-B (10)	10	<.0000014
Deaf-A (10)	9	<.0000413
Deaf-B (10)	10	<.0000014
Psychologists (10)	10	<.0000014
Total (50)	48	<.0000001

Note: Probabilities from Table 3, Feller (1950, p. 98). Exact probabilities computed by the binomial expansion. Group comparisons could be made on the basis of three or fewer correct matches ($p = > .05$) and four or more correct matches ($p = < .02$); this would reduce but not eliminate the high level of significance.

An analysis of variance produced an F of 5.48, significant at the .01 level, indicating an over-all difference among the five experimental groups. The mean matching accuracy of these groups was as follows:

Normal group A (average I.Q.)	4.3
----------------------------------	-----

Normal group B (superior I.Q.)	7.1
Deaf group A (average I.Q.)	3.0
Deaf group B (superior I.Q.)	4.9
Clinical Psychologists	7.0

Table II summarizes the significant differences between these groups.

As reported in the Table, neither of the deaf groups excelled in the number of correct matches. Likewise, though the clinical psychologists were relatively high in average matching ability, their performance was equaled by that of Normal Group B. Thus the two "expert" groups were not demonstrably superior in the number of correct matches.

Correlation coefficients were computed for each of the normal and deaf groups, between I.Q. and the number of correct matches. In Normal Groups A and B, matching accuracy was positively correlated with I.Q. (.63 and .65 respectively), significant at the .05 level. However, the correlations between I.Q. and matching accuracy for Deaf Groups A and B were $-.31$ and $-.35$. An analysis of covariance was next performed with the Normal and Deaf groups separately. With intelligence thus statistically controlled, the significant differences in mean matching accuracy between the two Normal Groups disappear. However, this was not true for the two Deaf Groups: with I.Q. statistically controlled, differences on the criterion variable remained significant ($p = < .01$). Finally, correlation coefficients between-groups and within-groups were computed from the data of the two covar-

TABLE II. Differences in Mean Matching Accuracy Between Each of Five Groups and Corresponding Levels of Significance

Group	Normal Group B	Deaf Group A	Deaf Group B	Clinical Psychologists
Normal Group A	2.80*	1.30	0.60	2.70*
Normal Group B		4.10**	2.20*	0.10
Deaf Group A			1.90	4.00**
Deaf Group B				2.10

*Difference in means significant at .05 level (2.16)

**Difference in means significant at .01 level (2.88)

iance analyses. Results again confirmed a significant tendency for normal subjects who were high in I.Q., also to be high in matching accuracy. Nonsignificant negative within-group correlation coefficients for the deaf groups indicated that a close relationship between I.Q. and matching accuracy did not obtain. That is, general intellectual level was positively correlated with matching accuracy within and between the two normal subgroups, and between, but not within, the two deaf groups.

All five of the experimental groups demonstrated a trend toward greater accuracy when a feeling of confidence in the matching existed. There were 190 matchings in each of the Most Confident and Least Confident categories (one subject failed to comply with the instructions): 107 matchings in the first category were in agreement with Steig, and only 53 in the second. However, when the nonparametric sign test was applied to determine the relationship of confidence to matching accuracy in each of the five experimental groups, only in the Deaf Group B did this relationship reach significance ($< .01$).

DISCUSSION

In general, the results of this study suggest that Steig did effectively communicate the intended meaning of some of his cartoons to fairly divergent groups. However, though all groups demonstrated matching accuracy far above chance expectancy, there were significant variations between some of the experimental groups, as well as within group variability. An analysis of the specific nature of the subjects' responses limits enthusiasm for the impressive statistical results.

For one thing, the average number of correct matchings for the two normal groups appeared somewhat related to the I.Q. averages of these groups. In the deaf groups, matching accuracy was not significantly correlated with I.Q.; however, it is possible that the

use of elaborated captions and an interpreter to assure an adequate comprehension of the captions may have been so successful as to obliterate individual differences in I.Q. within these groups, though Group B (the more intelligent) remained significantly better in performance than Group A.

Although Machover has suggested that the meanings of these figure drawings can be understood spontaneously and directly, she has also declared that ability to interpret such drawings meaningfully can be improved with "proper training and experience."

As with all projective techniques which seek to grasp as a whole and yet analyze the complex pattern of personality organization, the technical psychological knowledge and the clinical sagacity of the analyst are indispensable to the translation of the language of the method into the traits and the behavioral dynamics and trends of the individual studied (*ibid.*, p. 104).

Thus it was predicted that clinical psychologists as a group, having had special training and experience in personality evaluation, would demonstrate a distinctly superior ability to perceive the intended meanings in Steig's cartoons. In a somewhat similar vein, the prediction of a superiority among the deaf in perceiving cartoon meanings seemed justified because of their particular dependence upon non-verbal cues for an understanding of the meanings and intent of others. Neither of these predictions proved correct however; while both groups demonstrated an ability to match cartoons and captions which far exceeded chance expectations, they were not demonstrably superior to the normal groups.

Another limitation relates to Machover's expectation that the meaning attributed to each of the cartoons would be communicated to the majority of subjects. It is interesting to note that over 50 per cent of the matching accuracy of the subjects in this study (132 out of a total of 263 correct

matches) was contributed by only six of the total 32 cartoons! In addition, there were four cartoons which, while not communicating the meaning intended by Steig, had a meaning upon which a considerable number of subjects agreed. For the remaining 22 cartoons, there was relatively little group agreement, either between the subjects or with Steig. For example, one cartoon was matched with 26 different captions!

Still another and related feature of this study deserves rather extended comment. To the majority of the 36 cartoons cited in her article, Machover gave colorful, dynamic descriptions. These descriptions which she maintained were congruent with the meanings intended by Steig, were rich in psychoanalytic and body image flavor, and redolent of the figure drawing interpretations reported in her test manual. Examination of these dynamic descriptions reveals Machover to be an individual with an acute sensitivity to the expressive forms of the drawings, and the ability to communicate this responsiveness verbally. When an interpretation given by Machover is compared with the cartoon itself, in virtually every case, the meaning she attributed to it seems plausible. However, most of the cartoons to which she gave credible interpretations failed to suggest the actual captions with comparable consistency in the matching procedure of this study.

An example of the lack of congruence in meaning is the cartoon captioned "Serene Feeling" by Steig. The cartoon itself presents a man completely separated from his hat, his pipe, his hands, and his feet. Machover's interpretation follows:

Man is serene only when he has removed his hands and feet, and thus can go nowhere nor contact anything — when he is turned aside, and when he cuts off the social appurtenances that support him (p. 453).

As one compares the cartoon with Machover's paraphrase of Steig's cap-

tion, the plausibility of the match must be readily admitted. An examination of the experimental data reveals, however, that only two of the 36 subjects were able to suggest this appropriate meaning when the caption was not actually appended to it. Of apparent greater potency as appropriate captions for this cartoon were the following: "All strung together" (three matches); "Jittery" (three matches); "I recreated myself" (seven matches); "No one is exactly like anyone else" (three matches); "Affable man" (four matches); and "My trouble is purely organic" (seven matches).

It is interesting to note the consistent persuasiveness of her various descriptions, all of which suggest an underlying certainty in attributed meaning. However, as was noted above, Machover's conviction or self-confidence about the interpretative meaning of the cartoons (which may be correct), was simply not reflected in consistent matching accuracy by the experimental subjects. And with a single exception (Dent Group B), there was no significant relationship between the subject's confidence in his matchings and their objective correctness. Thus, there appeared to be no consistent relationship between the "just right feeling" experienced by Machover and the subjects as regards the communication of the intended meanings in these particular cartoons.

Such results cast doubt upon Machover's assertion that the "meanings" of Steig's drawings are spontaneously and universally understood, even by a group presumably possessing the "technical psychological knowledge and clinical sagacity" cited as facilitating refined analyses of figure drawings. These findings are reminiscent of Fisher and Fisher (1950) who studied inter-judge agreement in the interpretation of figure drawings using psychologists, psychiatrists, and stenographers. They found poor agreement among judges, and furthermore that persons untrained in formal psychol-

ogy were no less acute than the other judges in their appraisals.

The continued popularity of the DAP in the absence of supportive research results requires some explanation. There is a combination of plausible factors, beginning with the fact that the test is exceedingly simple to administer. Another is that in the absence of empirical standards some clinicians may find comfort in interpretations free from the threat of contradiction. Somewhat related is the observation that any ambiguous projective product invariably confounds "projection" and "identification" on the part of the examiner; human figure drawings seem a particularly potent means of eliciting projections from the clinician, which are mistaken by him for empathy and intuition. The fact that the DAP is usually administered within a battery of psychological tests contributes to the possibility that meaning will be "read into" the Test and that the resulting experience of success of interpretation is spurious. Finally, there is the psychologist's faith in projective methods of personality assessment and the belief that figure drawings *must* be dynamically meaningful projections of personality. There is a rich and provocative body of theoretical literature to support this supposition. In fairness, it must be emphasized that most of these objections are applicable to all projective methods, of course.

Suggestions for future research

Steig's own captions were utilized in this study as verbal equivalents to the meanings assumedly communicated by the cartoons because of their directness and succinctness; however, it is possible that these captions do not represent adequate equivalents to the meanings of the drawings. One alternative would be to use a broad spectrum of generic descriptive terms (e.g., love, happiness, fear, contempt, etc.), and ask subjects to sort cartoons into the categories they felt best approximated their meaning. This procedure

might markedly reduce intersubject variability because the obvious overlap in meaning between some of the cartoons would be taken into account. The general feeling tone of the cartoon would thus be gauged rather than the relatively specific and narrow meaning usually designated by Steig's captions. Osgood's semantic differential offers the possibility of quantifying such an approach. Perhaps the appropriateness of Steig's captions could be determined by having both drawings and caption rated on the differential.

Another possibility would be to eliminate Steig's captions and substitute Machover's rich descriptive interpretations as the meaning equivalent—this would constitute a more direct evaluation of her hypotheses. Or captions might be eliminated entirely by the process of having subjects group the drawings that "belong together." This would control the factor of verbal comprehension. No earlier study was found in which the factor of intelligence played a role of such importance in the judgment of expressive behavior. It is entirely possible that the nature of the verbal equivalents (captions) utilized in this study exaggerated this relationship.

The hypothesis that the congenitally deaf should have an enhanced sensitivity to "body language" remains plausible and a retest using an altered matching assignment could highlight this ability. It would also be of interest to use other special subject groups, such as paranoid schizophrenics, artists or actors, normals differing on measured social adjustment and social sensitivity, or Witkin's "field dependent" and "field independent" personalities, to cite a few examples (Moss 1957, Wolff 1943, Witkin 1954).

A final variation relates to the cartoons themselves. There is the possibility of using judges to grade the pictures along a continuum of directness or literalness and abstractions or symbolism. Results of this investigation suggest that the literal drawings com-

The Role of the TAT in the Measurement of Achievement as a Function of Expectancy

BERNARD I. MURSTEIN
Interfaith Counseling Center
and

HERBERT L. COLLIER
University of Portland

The achievement motive has been the subject of a large amount of research by McClelland (McClelland, Atkinson, Clark, and Lowell, 1953), Atkinson (1958) and their associates. In his book Atkinson presents a theory of motivation to account for the measurement of motives via the Thematic Apperception Test. The theory holds that the TAT response is a function of the product of the motives operating in the S and his expectancies of goal-attainment derived from situational clues at the time of the testing situation. Expectancy may be further broken down into two components; (a) subjective probability of success P_s and (b) incentive $(1-P_s)$. The use of the term $(1-P_s)$ as the measure of incentive follows the reasoning that the goal which is the least likely to be attained possesses the greatest attraction. The approach gradient of the strength of motivation to achieve as measured by a thematic test is represented as follows:

$$\text{Motivation}_{\text{ach}} = \text{Motive strength}_{\text{ach}} \times P_{s_{\text{ach}}} \times (1 - P_{s_{\text{ach}}})$$

where P_s represents the subjective probability of success.

Where data are available regarding S_s "fear of failure", the avoidance gradient may be obtained using the formula:

$$\text{Motivation}_{\text{avoid failure}} = \text{Motive Strength}_{\text{avoid failure}} \times P_{s_{\text{avoid failure}}} \times 1 - P_{s_{\text{avoid failure}}}$$

The resultant motivation would be equal to the motivation to achieve minus the motivation to avoid failure.

P_s may be differentiated from objective probability of success (P_o) since a person's confidence in his ability to beat the objective odds in a risk-taking situation will not necessarily follow the objective probability of success for the group as a whole. Suppose that a group of 20 students are told that the top ten in an arithmetic test will receive a prize. If this is a random class, there will probably be some individuals who think highly enough of their ability to assess their own chances as greater than .5. Other individuals may believe they are below the mean in ability and may assess their chances as considerably less than .5. In the absence of any measure of subjective probability, P_o is used as a substitute for the median P_s for a group in our study. We assume that the group is a representative one with regard to the self-evaluation of success for the task it is assigned. In a representative group, therefore, for some persons $P_o > P_s$ and for others P_o is .5 and $(.5) \times (1-.5) = .25$. The median P_s .

If we now hold motive strength constant, the motivation to produce a projective response reflecting achievement is a function of $P_{o_{\text{ach}}} \times (1 - P_{o_{\text{ach}}})$. It may readily be seen that this product is at a maximum when P_o is .5 and $(.5) \times (1 - .5) = .25$. The product of probability \times incentive is minimal when P_o is either very high or very low. Thus, when $P_o = .9$, $(.9) \times (1-.9) = .09$.

The purpose of this experiment was to test the validity of the Atkinson theory by holding motive strength constant via randomization and vary-

ing P_o in a series of tests (arithmetic, cancelling) similar to that employed by Atkinson (1958 b). The motivation to avoid failure was unknown but was assumed to be approximately equal for each group due to the use of randomization in the selection of Ss.

The hypotheses were: The $P_o = .5$ group will project significantly more achievement fantasy on a series of thematic cards than a group where (a) $P_o = .1$, and (b) another group where $P_o = .9$.

PROCEDURE

Sixty children (30 boys and 30 girls) from the seventh grade of a Portland public school were randomly selected from a larger group of 150 students in the same grade. These 60 children were randomly assigned to one of three subgroups each containing 10 boys and 10 girls. The purpose of the randomization was to equate the groups for n-achievement motive strength. Each of the groups was tested separately on the same afternoon. The students were instructed that they would receive a series of tasks in accordance with a broad study of the seventh grade students.

In an attempt to eliminate the idea that intellect or skill played an important role, Ss were further instructed that the harder they worked the better their score would be.

A conference with the seventh grade instructors over the choice of rewards for successful performance on the tasks led to opposing views. One group felt that a material reward would most motivate the group. The other group held that a "flashy" certificate would yield the greatest effort on the part of the Ss. Accordingly, it was decided to incorporate both features in the reward. Each child who finished above the cutting point for his group would receive a card with a shiny gold seal attached to the card. The card stated

"Pay to the bearer one ice-cream cup". E assured the Ss that each winner would definitely receive a certificate and ice-cream cup when the results of the tasks had been tabulated on the day following the tests.

The instructions to the three groups diverged at this point. Group 1 was told that the top 2 out of 20 in the group would receive the prize. Groups 2 and 3 learned that the top 10 and 18 respectively would receive the prize. P_o of success therefore, for the three groups was .1, .5 and .9 respectively.

Following the statement regarding the probability of success, the Ss were told that before starting on the actual tasks, E would like to have them write stories for a series of slides which would be flashed on the screen. At this point the Ss were given booklets which contained the following questions for each slide:

1. What is happening? Who are the persons?
2. What has led up to this situation? What has happened in the past?
3. What is being thought? What is wanted? By whom?
4. What will happen? What will be done?

The slides used were, in order, TAT cards 1, 2, 6BM, 8BM, and Symonds Picture Story cards A-1 and B-6. These cards were chosen on the basis of apparent stimulus pull for achievement themes. Ss were told that each picture would be flashed on the screen for 20 seconds after which it would be withdrawn and Ss would have 4 minutes allotted to write their stories.

Following the showing of the slides, E presented the first of the two tasks. This task consisted of relatively simple arithmetic addition and subtraction problems. An example of one such problem is $43 - 15 = ?$ S was stopped at the end of four minutes and then given the second task which consisted of crossing out as many circles as possible from several pages containing

many rows of circles. At the end of four minutes, *E* again stopped *S*.

RESULTS

The thematic stories were scored for achievement according to the method described by Atkinson (1958), while the number of items attempted constituted the arithmetic and the cancelling scores. The stories were scored by the second author who established his scoring reliability as .86 using the achievement manual appearing in Atkinson's book (1958) as the criterion. The means and standard deviations of the fantasy scores for the three groups are shown in Table I. It is apparent by inspection that no significant difference exists between any two of the groups. The hypotheses regarding the superiority of the $Po = .5$ group to both the $Po = .9$ and $Po = .1$ groups are therefore rejected.

TABLE I. Fantasy Achievement Motive Means and Standard Deviation Scores of the Three Groups

	Group 1 $Po = .1$	Group 2 $Po = .5$	Group 3 $Po = .9$
Mean	8.50	8.25	8.55
Standard Deviation	3.41	2.23	3.66

A question that might be raised at this point is whether there were any differences among the groups in the arithmetic, cancelling, or combined tasks. Table II shows the means, stand-

ard deviations and *t* tests for both tasks individually and for the average of two tasks after conversion to T scores.

It may be seen that the $Po = .5$ group was significantly higher than the other groups on the arithmetic task but no significant differences were noted for the cancelling task. The combined task T scores also revealed no noteworthy differences. A nagging thought concerned the possibility that for our relatively unselected group, the arithmetic scores might be a function of intelligence and that possibly the $Po = .5$ group had by chance been assigned more intelligent students than the other groups. The arithmetic task, therefore, was correlated with the California Mental Maturity IQ-equivalent scores obtained for 57 of the 60 children. The resulting *r* of .23 just reached the significance at the .05 point for a one-tailed test. The relationship between these variables while significant, accounts for only a rather minute 5 percent of their total variance. Moreover, the mean IQ-equivalent scores for the $Po = .1$, $Po = .5$, and $Po = .9$ groups were found to be 107.9, 104.3, and 109.8 respectively. None of the differences between any two groups reached significance at the .05 level. It is apparent therefore, that intelligence does not account for the differences between the groups on the arithmetic task.

TABLE II. Means, Standard Deviations and *t* Values Between Three Expectancy Groups for Arithmetic, Cancelling and Combined Task Scores

Task	Probability Groups	Mean	S.D.	Groups Compared	<i>t</i>
Arithmetic	$Po = .1$	23.8	8.8	$Po.1$ vs. $Po.5$	3.60**
	$Po = .5$	30.6	14.0	$Po.1$ vs. $Po.9$	N.S.
	$Po = .9$	26.0	8.8	$Po.5$ vs. $Po.9$	2.43*
Cancelling	$Po = .1$	291.0	31.0	$Po.1$ vs. $Po.5$	N.S.
	$Po = .5$	288.5	50.6	$Po.1$ vs. $Po.9$	N.S.
	$Po = .9$	296.7	39.3	$Po.5$ vs. $Po.9$	N.S.
Combined Tasks (T scores)	$Po = .1$	49.9	9.5	$Po.1$ vs. $Po.5$	N.S.
	$Po = .5$	49.7	8.0	$Po.1$ vs. $Po.9$	N.S.
	$Po = .9$	48.7	8.2	$Po.5$ vs. $Po.9$	N.S.

* significant at .05 level

** significant at .01 level

Granting that the expectancy of reward did not affect the manifestation of the achievement motive via fantasy, the question remains as to whether Ss distinguished by high vs. low achievement fantasy scores differed significantly in achievement on their arithmetic or cancelling scores. Three chi square analyses were done pitting fantasy scores against the arithmetic, cancelling and combined task scores in turn, with the median serving as a cutting point. The analyses yielded values of .12, .22, and .00, all clearly non-significant.

To hold the differing probability values constant the same analysis also was computed within each probability condition for each task, but again no significant chi square values resulted.

An overall correlation between fantasy scores after conversion to T scores, and the arithmetic, cancelling, and combined scores yielded non-significant values for the total group of 60 Ss as indicated in Table III. This table, however, indicates the significant correlations for Group 3 ($P_o = .9$) between the fantasy and cancelling scores ($p < .05$) and the fantasy and combined scores ($p < .01$). It seems reasonable to conclude, therefore, that persons high on the achievement motive as demonstrated by their arithmetic and cancelling scores could not be predicted by their achievement fantasy scores without knowledge of the probability group from which they stemmed.

Finally, the cards were inspected for "stimulus pull" by noting the number of achievement themes that each card

yielded disregarding the various probability subgroups. The results were as follows: TAT card 1, 24; card 2, 12; card 6BM, 2; card 8BM, 3; Symonds card A-1, 0; card B-6, 2. Thus, the cards differed greatly in their ability to elicit achievement themes.

DISCUSSION

In a limited way our data supports Atkinson's view that motivation is strongest under the condition of maximum uncertainty ($P_o = .5$) in that this group attempted more arithmetic problems than either of the other two groups. No significant differences, however, were found for the cancelling or combined task scores.

The role of fantasy as an index of motivation to achieve seems questioned by our results both because it failed to differentiate the three probability groups and because it failed to correlate with achievement in the various tasks for the pooled Ss.

The finding that the $P_o = .9$ group showed positive correlations between achievement fantasy and each of the three performance scores merits some discussion. "Fear of failure" may be presumed to have a deleterious effect on efficiency of performance. (Atkinson and Litwin, 1960). Also, "fear of failure" may have been maximally aroused for the $P_o = .9$ group because failure represented a "loss of face" when 18 of 20 succeeded. It might be expected then that the more anxious person in whom "fear of failure" was high would manifest a consistently low score on both the performance and fantasy tasks. Those

TABLE III. Correlations Between Fantasy and Arithmetic, Cancelling and Combined Task Scores

	Fantasy vs. Arithmetic	Fantasy vs. Cancelling	Fantasy vs. Combined Task
Total Group	-.13	.14	.23
Group 1 ($P_o = .1$)	-.04	-.09	-.01
Group 2 ($P_o = .5$)	-.09	-.14	-.06
Group 3 ($P_o = .9$)	.14	.43*	.57**

* Significant at .05 level for a two-tailed test

** Significant at .01 level for a two-tailed test

persons in whom "fear of failure" was low would perform fairly high on both tasks. We are thus speculating that we accidentally selected a bimodal sample for "fear of failure" through our randomization procedure.

It should be noted that Atkinson and Litwin (1960) have shown that "fear of failure" and the achievement motive are only slightly associated ($r = -.15$). Thus, we may presume that while the approach gradient of the achievement motive was most strongly elicited in the maximum uncertainty group ($P_o = .5$) the "fear of failure" motive was not. Hence, the correlation between fantasy and performance should be at a maximum only when the "fear of failure" motive is strongly aroused. The reader will recognize the highly speculative nature of this argument and we are inclined to await replication of the study rather than pressing the issue.

Since our data do not generally support Atkinson's theory, it should be noted that our procedure differed considerably from that usually followed by Atkinson and his associates. First, our group is far more representative of the general population than is the select group of college students employed in the majority of their studies. Secondly, our sample consisted of 12 and 13 year old children rather than college students. Third, we employed both boys and girls while they used only men in the major portion of their research. Fourth, the stimulus cards used by them were not the same as those we employed. It was necessary for us to choose cards whose stimulus pull was assumed to be maximal for children, rather than for adults. In this regard, we achieved only partial success. Only two of our cards (TAT 1 and TAT 2) produced any substantial number of achievement themes. It is possible that if we had used cards more highly structured for achievement, a greater spread might have been obtained between our groups. Supporting this view are the results

of Kagan who found cards highly structured for hostility to be more effective than ambiguous cards in differentiating hostile and non-hostile boys.

Still another factor possibly contributing to the differences in results is due to the time limitation. It was possible to work with the children only 45 minutes. Accordingly, the time limits of 20 minutes for the arithmetic tasks and for the cancelling tasks, set by Atkinson and Reitman (1958), were trimmed down to four minutes for each. Nevertheless, significant differences were obtained for the arithmetic task despite the shorter time allowance. The cancelling task did not reveal significant differences however, which may have been due in part to the time element.

A further factor to be considered is the possibility that the experimental situation was itself so stimulating to our students that they were striving to achieve to their utmost regardless of the probability of achieving a prize. This interpretation would find support in the results of Atkinson (1958 b) who showed that differences between high and low achievers tended to decline for all probability situations when the prize was increased from \$1.25 to \$2.50. The fact that the $P_o = .5$ group did significantly better than the other groups on the arithmetic task however, seems to negate this possibility.

In sum therefore, the function fulfilled by fantasy expressed on the TAT seems to still remain somewhat of an enigma. It is hoped in future research to deal more thoroughly with the problem of the optimum gradient of stimulus structure to employ in making the thematic cards maximally sensitive to differential states of motivation. It would seem that the inclusion of thematic cards whose stimulus properties are not precisely known with respect to the various need-dimensions (e.g., achievement, affiliation), makes it difficult to adequately test any

theory relating to the nature of fantasy. (Murstein, 1960).

SUMMARY

Sixty students (30 boys and 30 girls) were randomly selected from all seventh grade students at a Portland elementary school. Ss were randomly assigned to one of three groups, the only stipulation being that the group be equated for sex.

All groups were separately told that high performance on a combination of a simple arithmetic task and a cancelling task would be rewarded by a prize of an ice-cream cup as well as a certificate of merit. The first group was told that the top 10 per cent would receive the prize, the second that the top 50 per cent would gain the prize, and the third that the highest 90 per cent would receive the reward.

Just prior to these tasks, 6 thematic cards (TAT 1, 2, 6BM, 8BM, and Symonds A-1, B-6) were presented to the Ss via slide projection and they were asked to make up stories to these cards. The stories were scored for achievement imagery by the method devised by McClelland and his associates. The hypothesis stemming from Atkinson's theory of motivation was that the group with the maximum uncertainty (whose objective probability of reward (P_o) was equal to .5) would be the most highly motivated and therefore manifest the most

achievement fantasy. This hypothesis was not confirmed. In view of the fact that the $P_o = .5$ group did do significantly better in the arithmetic task than the other groups, the results are interpreted as detrimental to the view that fantasy directly reflects the achievement motive. Possible causes for the differences between the findings of this study and earlier studies reported by Atkinson are discussed.

REFERENCES

- Atkinson, J. W. *Motives in fantasy, action and society*. Princeton: D. Van Nostrand, 1958 (a).
- Atkinson, J. W. Towards experimental analysis of human motivation in terms of motives, expectancies, and incentives. In Atkinson, J. W. (Ed.) *Motives in fantasy, action and society*. Princeton: D. Van Nostrand, 1958 (b) Pp. 288-305.
- Atkinson, J. W. & Litwin, C. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *J. abnorm. soc. Psychol.*, 1960, 60, 52-63.
- Atkinson, J. W., & Reitman, W. R. Performance as a function of motive strength and expectancy of goal attainment. In Atkinson, J. W. (ed.) *Motives in fantasy, action and society*. Princeton: D. Van Nostrand, 1958, Pp. 278-287.
- Kagan, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, 52, 390-393.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
- Murstein, B. I. The measurement of ambiguity in thematic tests. *J. proj. Tech.*, 1960, 24, 419-423.

Received June 13, 1961

Projective Techniques in the Selection of Management Personnel

JOSEPH G. PHELAN
Los Angeles State College

Studies of management selection in industry and business have generally set forth the virtues of one test or technique or another. Concentration has been on the technical aspects of the jobs, including lists of duties, responsibilities, functions and activities performed. Some emphasis has been placed on the personality traits of individuals filling particular jobs, on means of estimating the presence or absence of the traits in question, and on prediction of success in leadership situations. Meehl (1959) failed to find evidence of projective technique validity in such predictions.

Improvement in prediction will depend on finding means to demonstrate empirically successful prediction, then isolating the factors which made prediction possible. Factors once pinpointed, must be operationally defined and established as valid.

Ghiselli (1955) insists that job analysis should precede research and use of methods to predict successful performance in management jobs. For the management series of positions, information on which to build job analysis is not forthcoming. Jobs differ greatly, one from another. Proficiency or job success measures are likewise not likely to be attainable in the near future.

At the present stage of progress of research in this area, the best approach seems to be to use the predictors at hand, then to evaluate their effectiveness by comparing predictions as to promotability to management positions with the subsequent records of the subjects as to actual promotion and ultimate job success.

The present investigation is concerned with the relationship between performance of a battery of predictors as to promotion to administrative po-

sitions and the ultimate promotion and success in promotion of the men in question. Such a study establishes validity of the selection procedure as a result of comparing the testors' predictions as to promotability of the candidates with their actual history of promotion to administrative posts. At the same time, the predictive efficiency of individual tests as to promotion can be made. Comparison of individual tests can be made with the combined integrated results of the tests used together.

PROCEDURE

94 industrial men were tested in groups of from 3 to 7 from 18 industrial organizations. Ages ranged from 18 to 43. Educations ranged somewhat; all were high school graduates, most were college graduates, liberal arts, accountants and engineers. Some were hourly workers, most were on salary. None were occupying administrative positions at the time of the study. All were being tested for first-line supervisory positions. Industries from which they came were in the eastern and midwestern sections of the United States. Products manufactured by the various companies included food, drugs, paper-products, petroleum, metal and wood products.

The top management in each industrial organization asked that we rate these men in terms of promotability to management positions. In time one man of each group was eventually promoted to a management position.

The tests taken by each of the management candidates included the American Council on Education Test of Scholastic Aptitude, the Beckman Revision of the Allport Ascendancy-Submission Test, the Bennett Test of Mechanical Comprehension, The File-

Remmers "How Supervise" test, the Rorschach and the Thematic Apperception Test.

Upon completion of testing, written reports on each man were submitted to the industrial organization which sponsored his testing. These reports varied in length from 5 to 7 pages. Conclusions on the personality patterns of the men as they applied to management jobs were based on information gleaned from the various tests used. This method of making predictions is considered to be holistic in the sense that the entire test record was analyzed and the report based on the overall impression.

The reports were written to deal with the following areas:

Relations Established with Others:

- Methods of dealing with others
- Ability to work with others
- Ability and preferences for working alone
- Ability to communicate orally and in writing
- Personal flexibility and adaptability

Reactions to supervision and Company Policy

Work Habits and Attitudes

- Industry and initiative
- Drive and energy
- Persistence and follow through
- Co-operation and friendliness
- Objectivity

Mental ability and use

- General level of intelligence
- Creativeness, originality
- Practicality, concrete mindedness
- Breadth of interests

After a period of not less than 18 months and not more than 3 years from the time that the reports were submitted to the various industrial organizations, representatives of the executive branch of the client companies were asked to rank each of the men who had been tested in terms of promotion or promotability. At the same time, two experienced personnel men, who had no part in the original

testing and evaluations and no connection with the industrial organizations from which these men came, were asked to read only the written reports and to rank each man of each company in terms of promotability to management positions. Using the Spearman-Brown rank order formula, the rankings of the personnel men were correlated with the companies' rankings.

RESULTS

The median rho, (rank-order correlation) for all 18 industrial populations was .70. With seven degrees of freedom, this relationship is significant at the .05 level. Of 18 rhos, eight are significant at the .05 level or better. These results exceed chance expectations. For three of the 18 companies, a perfect correlation was obtained; for an additional three, rhos were at .90 ($p > .01$) and for another group of three companies rank order correlations were at .80 ($p > .02$).

Of the 18 groups who came to the laboratory for appraisal the best man, that is the man who was subsequently promoted or who was considered by the company to be most promotable was also so selected by the independently rating personnel men in 12 of 18 cases. This 66% selection of the top candidate is statistically significant. If a one step interval is permitted, that is to say if we include selections where the personnel man picked a man as top who turned out to be rated second best by the industry, matching performance is at the 83% level of correctness ($p > .01$).

In addition to evaluating the predictive efficiency of the overall test battery, the predictive power of individual tests of the battery was examined.

For 18 companies, the American Council on Education Examination yielded a median rank order correlation of .30; the Bennett Mechanical Test rho was = .21; The A-S Reaction Study rho = .30; How Supervise rho = .20. None of these is significant.

In addition, a trained psychologist interpreted the complete Rorschach protocol for each man and the T.A.T. material. On the basis of these data, singly and combined, he ranked the same 94 applicants, by company in terms of promotability. The psychologist's rankings were compared with company rankings. The Rorschach alone yielded a median rank order correlation of .30 which was not significant; T.A.T. correlation was .41. When the judgments as to promotability were based on Rorschach and T.A.T. interpretations, used in combination, the median rho was .56 with $p = .10$. Testers reported that the projective tests, in this case the Rorschach and the Thematic Apperception Test, seemed to form the nucleus of the pattern which enabled the predictors to perform at a high level of predictive efficiency. T.A.T. and Rorschach, used together yielded a higher correlation than any objective test used singly. When psychologists were asked which tests they found most useful in making predictions, they indicated a preference for the projective techniques.

SUMMARY AND CONCLUSIONS

Although management jobs vary, to a degree, from company to company so that it is difficult to specify the precise nature of the job for which the candidate is considered, predictions of a high degree of success as to fitness for administrative jobs are possible over a wide range of roughly similar management positions.

The integrative, or holistic method of interpreting test results yielded

higher correlations with the criterion, prediction of promotability to administrative positions than did the scores of individual objective tests.

In general, individual objective tests have little predictive value when considered separately.

Projective tests taken singly, do not predict at a highly significant level. Rorschach and T.A.T., taken together, were the best single predictor.

When a battery of objective and projective tests are used together and all clues and scores are integrated into a pattern by trained analysts, valid predictions as to success in industrial supervisory jobs can be made.

In practice, supervisors are not evaluated with a single measuring instrument or technique. Often the instruments used do intercorrelate to a degree, but measure slightly different aspects of the same ability.

Intensive integrative analysis of all test responses, when done by trained psychologists, yields better results than do numerical test scores.

When the integrative approach was used, the projective tests, Rorschach and T.A.T. provided most of the information on which the successful prediction was based.

REFERENCES

- Ghiselli, E. E. & Brown, C. W. *Personnel and industrial psychology*. N.Y.: McGraw-Hill, 1955.
- Meehl, P. *Clinical versus statistical prediction*. Minneapolis: U. Minn. Press, 1954.

Received April 24, 1961

Revision received August 1, 1961

Intraindividual Response Variability in Figure Drawing Tasks^{1, 2}

JOEL SHANAN

Department of Psychiatry, Hadassah University Hospital, Jerusalem

The purpose of this study was to investigate the effects of different stimulus conditions on intraindividual variability of one specific response to the Draw A Person Test (DAP) (Machover, 1949), namely the size of the drawings. Attention was focussed on some of the theoretical notions of Fiske and his co-workers concerning the problem of intraindividual variability in test responses. Fiske (1957) states: "When the intensities of external stimuli are low an organism shows marked variability of behaviour, presumably because of the ever changing neural and physiological states within the organism. When external stimuli are at moderate levels of intensity the extent of variability is reduced; responses show some consistency. Thus internal stimuli may be associated with variability. Some external stimuli and conditions are more constraining than others. The greater the total influence of such constraints the less the variability of behaviour."

Fiske considers items, instructions and other aspects of the testing situation as such constraints on the potential variability of response. If the item is held constant, test instructions may be classified in terms of their stimulus strength or intensity and consequently in terms of their potential power to constrain response variability.

One of the possible criteria for such a classification may be the amount of ambiguity of the stimulus presented in the instructions (Fiske, 1957).

When a verbal stimulus is worded clearly, i.e. is more structured, it may be considered as strong. On the other hand verbal stimuli lacking clarity because of their wording novelty or other reasons can be considered as weaker. Instructions given on repeated occasions would represent more structured stimuli than when given the first time. Internal psychological stimuli such as attitudinal or perceptual sets, or anxiety attached to such sets can also be assumed to influence response variability according to their intensity. While this assumption is at variance with Fiske's (1957) earlier views³ on the effects of internal stimulation, it follows the logic of his argument concerning the relationship between stimulus intensity and response variability.

The present study, was mainly concerned with testing within the context of a conventional testing situation, the hypothesis derived from Fiske's theorizing that conditions which raise the degree of ambiguity of the task—at least within the realistic limits of the testing situation—will be reflected in a rise of intraindividual response variability. As mentioned, size of drawing⁴ was chosen as the dependent variable and interest was focused specifically on the question of how the various testing conditions would affect intraindividual variability of this response. According to our general hypothesis it was expected that figure

³From a personal communication by Dr. Fiske it appears that he now believes that spontaneous expressive tendencies rather than internal stimulation per se, raise response variability.

⁴Figure size was chosen for the following reasons: a. It is easily and reliably measurable. b. Because of the importance and relative stability attributed to it by the experts of the "draw a person" test. (Levy, 1952, Machover, 1949, Machover 1951).

¹This study was supported in part by a grant from the Hadassah Medical Organisation.

²The author is grateful to Dr. D. W. Fiske for his encouragement and for critical reading of and helpful comments on earlier drafts of this paper.

size in the DAP Test would show more consistency, when,

1. Instructions were relatively clear, (such as the request "draw a person of the opposite sex"⁵ or a person in a specific social role, such as asked for in the instruction "draw a teacher") as compared with drawings performed under more ambiguous instructions (i.e. "draw a person").
2. When performance was repeated under similar conditions.
3. When the subjective definition of the ambiguous task was based on a relatively clear internal perception.
4. When two or more constraints functioned at the same time.

METHODS

The subjects of this study were 54 Jewish girls between the age of 18-22. All of them were of middle-class background and had completed high school. Most of them had had one or two years of college education. In terms of their country of origin there were 20 from Latin American countries, 18 from Anglo-Saxon countries and 16 of Israeli origin. While the study was carried out all subjects were living in co-educational boarding schools, where their daily activities were organised in a fairly uniform way. The stay in these boarding schools represented for all subjects their first prolonged separation from their parental home.

Participation in the study was on a voluntary basis though some pressure to participate was exerted on the part of the schools. S's were told that they were to participate in a study of adjustment to new life conditions. A series of three figure drawing tasks

was administered in two group sessions. The first session was held shortly after the arrival of the subjects in the boarding school and the second session six months later. S's were addressed in their native language by proctors in the presence of the author. Wording and sequence of instructions were identical in both sessions: First the group was asked to "draw a person (a woman or a man—a man or a woman as you like)"⁶. After completion of the first task subjects were asked to "draw now a person of the opposite sex". Finally foreign subjects were asked to "draw a pioneer girl" and Israeli subjects to "draw a teacher".⁷

The area of each drawing was measured with a planimeter and its size was measured in millimeter square. Accessories such as tools, furniture, etc., or objects other than the person were not included in the measurements; however any item representing clothing was. Open spaces such as between arms and body or the space between the legs of the figure were not included. Each figure was measured twice. Reliability of the measure was close to one, but for final computations the arithmetical average of the two measurements was used. Spearman rhos were used to measure consistency—the complement of intraindividual variability of responses—within sessions (i.e. the relationship between the different drawings drawn in one session) and between sessions (i.e. the relationship between the first drawings in sessions one and two, etc.)⁸

⁵This departure from standard instructions was necessary because of the difficulty of translating "person" into Hebrew and Spanish without implying sex of figure.

⁷These two roles were chosen because they represented educational goals for the respective groups of respondents.

⁸This question is discussed at greater length in a separate communication dealing with the clinical implications of intraindividual variability in figure drawing tasks (Shanan, 1960.)

⁶The instruction "draw a person of the opposite sex" was presented according to original test instructions after the respondent had completed the first drawing and thereafter the sex of the second drawing was determined by the sex of the first one.

RESULTS

In order to test whether clarity of instructions would constrain variability, size of drawings performed under more and less ambiguous instructions was compared. From Table I which shows the results of these comparisons, we learn that variability was lower under the clearer instructions as compared to the ambiguous instruction "draw a person". While the difference between the two types of instructions might not seem very impressive the finding itself appears to be fairly stable. As indicators of response consistency Rho's were computed for the three ethnic sub-groups independently. These rho's ranged from .21-.52 for responses under instructions "draw a person", from .34-.68 for responses under instruction "teacher", and from .45-.85 for responses under instruction "opposite sex". In all cases it could be seen that variability was highest, i.e. consistency was lowest under instructions "draw a person" as expected in terms of the hypothesis. It may also be added that in all subjects the difference between amount of variability under opposite sex instructions and amount of variability under DAP instructions was always greater than the difference in variability between "opposite sex" and "special role" (teacher, etc.) conditions. These data seem also to indicate that behavior was atypical under DAP instructions at the beginning of a session.

In order to find out whether repetition of task would constrain intra-individual variability and thus raise consistency, the respondent's consistency within one session over different instructions was compared with his consistency over sessions for identical tasks or instructions. Then the level of consistency in session 1 was compared with that in session 2, the latter being considered as a repetition of the whole series of tasks.

Comparing the data of Table II with those of Table I we find that repetition of task within one session reduces variability considerably. This in spite of the fact that the drawings within one session were carried out under different instructions while the data in Table I are derived from repetition of tasks under identical instructions within an interval of six months. Comparing the level of intraindividual variability within session 1 with that in session 2, a relatively small drop in overall variability might be noted from session to session. It may be added that a group of 22 subjects was administered a third session immediately after the second one. Rho's obtained on within-session comparisons of the third series of drawings were very similar to those obtained in the second session. An additional very slight drop of variability was noted though. These findings then tend to support the notion that repetition of task will lower variability. It appears as if a person be-

TABLE I. Intra-individual Variability in Figure Size from
Session I to Session II
Three Sets of Instructions (N = 54)

Instructions Rho	Draw a Person .38	Opposite Sex .50	Special Role .49
---------------------	----------------------	---------------------	---------------------

TABLE II. Intra-individual Variability in Response to Three Different but
Similar Tasks Within Sessions (N = 54)
Instructions

Session	Draw a Person Opposite Sex	Draw a Person Special Role	Opposite Sex Special Role
Session I	.61	.64	.71
Session II	.80	.72	.72

comes more consistent on a task after repeating it. It even appears that this tends to hold regardless of minor changes in instructions.

The expectation that internal set would affect response variability differentially was tested by comparing the size of figures drawn as females with that of figures drawn as males—all drawn under identical, i.e. "draw a person" instructions which enabled the respondent to define the task according to his own needs in either way. Finally, in order to test whether the effects of constraints are additive, the combined effort of interpretation of tasks (i.e. drawing of either male or female under ambiguous instructions) and instructions (ambiguous versus non-ambiguous) were investigated. Size of female figures drawn under DAP instructions in Session I was correlated with size of female figures drawn by the same subject under "opposite sex" instructions in Session II and vice versa. The same procedure was used on male drawings. In other words, individual pairs of all identical sex figures were correlated between sessions—except those drawn both times under identical instructions. Table III presents the results of these comparisons.

Table III shows that the subjective interpretation of the ambiguous instruction as to the sex of the figure influences consistency. Female figures are more variable in size than male figures and this holds for all conditions. Fiske's earlier notion (1957) then, that internal stimuli (if subjective interpretation of the instruction by the subject is considered as such an internal, psychological stimulus) raise variability cannot be supported in an unqualified way. As one

would expect, some such stimuli apparently raise variability more than others. It is noteworthy that consistency of female subjects is lower on a performance which is supposed to reflect then sex identification (Machover, 1949). This, however, was the case under all test conditions. Finally in support of the notion that constraints act additively it was found that variability was highest on female drawings under the ambiguous "draw a person" condition, showing the cumulative effect of two contemporaneous stimulus conditions.

DISCUSSION

While most of the empirical evidence for Fiske's earlier theorizing (Fiske and Rice, 1955) on intraindividual variability was provided from studies based on an inductive approach, only more recently direct attempts were made to test hypotheses derived from Fiske's theorizing (Fiske 1957 a, b, 1959, Fiske and van Buskirk, 1957, Fiske and Cox, 1952, Mitra, and Fiske, 1956, Osterweil and Fiske, 1956, van Der Veen and Fiske, 1960). The present study can be viewed as a further effort to control some conditions affecting intraindividual variability or consistency, its complement, in a way which permits a direct testing of hypotheses derived from Fiske's theory. In using completely reliable and objective measures of the dependent variable, influence of judgmental processes on variability was eliminated. Finally, instructions, time and indirectly, response variability, were manipulated in a situation which, while allowing for some desired controls, was essentially a usual test situation. Thus the painful necessity to validate in life situations generalisa-

TABLE III. Variability of Figure Size over Two Sessions as a Function of Instructions and Individual Interpretations of Instructions

Interpretation of instructions as:	Instructions DAP	Regardless of Instructions
	Rho	Rho
Female	.39	.48
Male	.50	.59

tions derived from "pure" laboratory settings was reduced.

On the whole, the present findings may be interpreted as giving some further support to the notion that external stimuli function at least within a certain range as constraints on response variability, i.e. raise consistency. Some support was also found for the notion that such stimuli constrain intraindividual variability differentially and do function additively. The findings on internal stimulation—if one accepts subjective interpretation of test instructions as an internal stimulus—do suggest that internal stimuli of this sort tend to affect response variability differentially as do external stimuli. They do not support Fiske's earlier notion that internal stimuli necessarily increase intraindividual variability⁹. This may be more true for biological internal stimuli than for psychological internal stimuli. Maybe other psychological stimuli than those used in this study, will have other effects. The present findings of course are partly dependent on the design used which allowed for separation of different internal stimuli yet did not allow for a strict separation of internal from external stimuli—partly because of the effort to keep the testing situation intact as much as possible.

Looking at the findings from the point of view of the amount of ambiguity in the various stimulus conditions, the evidence supports the notion that stimuli can be ordered in this way and that consistency on a task is dependent on the amount of ambiguity in the stimulus condition; the greater the ambiguity of the stimulus, the greater the amount of intraindividual variability of the response, i.e. the less consistently does the respondent react to the given stimulus. If two or more conditions act in the direction of greater ambiguity of the stimulus, such as was the case when the ambiguous "draw a person" in-

struction was interpreted as a request to draw a female¹⁰ or when novelty was interacting with other conditions, variability tends to rise, i.e. consistency is lower.

The differences in response variability under the various stimulus conditions, as mentioned earlier, were consistent though not very impressive from a quantitative point of view. It should be kept in mind that the size of the figure in the DAP test is considered by Machover a relatively stable response (Machover 1949). Similarly the behaviour presumably measured by this response is traditionally considered a relatively stable aspect of personality. The findings tend to show that the consistency of even presumably stable behaviour is subject to fluctuations, the extent and nature of which can be related to stimulus conditions in a given test situation. This finding parallels earlier findings which demonstrated the effect of test conditions on the variability of response content (Osterweil and Fiske, 1956). There is little doubt that information on the degree of intraindividual stability of test responses under given conditions may contribute to the predictive value of the response. This, as Fiske has already pointed out, may be of particular importance in the investigation of the predictive value of projective test responses (Fiske 1959) and may eventually contribute to the clarification of some issues in the controversy on clinical versus statistical prediction. A fuller discussion of the implications of the present findings for the clinical use of the "draw a person" test will be presented elsewhere (Shanan 1960).

In conclusion it seems tempting to

¹⁰According to a number of studies the female role is perceived less clearly in western culture even by female subjects than the male role (Lynn, D. B. 1959). The lower consistency found on the female drawings is not quite in line with some of the clinical assumptions on the test (Machover, K. 1949). This issue is discussed elsewhere (Shanan, J. 1960).

⁹ c.f. footnote 3.

ask the following questions: If there is any relationship between various degrees of consistency in response and the anxiety level operating in the respondent under given conditions? And if so, what could be the possible relationship between anxiety and intraindividual variability or individual response consistency? Looking at the results from this point of view it appears that variability is high under all conditions which might be interpreted as anxiety arousing. Response consistency rises or levels off under conditions which may be presumed less anxiety arousing. In other words, it appears that individuals tend to act less consistently — and atypically — as compared to subsequent behaviour in novel or ambiguous situations, and their behaviour tends to become more consistent once they adjust to the situation by sharpened perception and familiarity with the stimulus conditions. It seems that novelty of the situation as a whole has a major influence on consistency of behaviour. Fiske has pointed to the possibility that variability factors may reflect acquired modes of adaptation (Fiske 1957). In the light of the present findings variability appears to be—at least under some specific conditions a correlate of anxiety. It could therefore serve as an index of specific manifestations of anxiety. This interpretation is in line with some of Fiske's recent findings: variability correlated among others with MAS scores and with measures of anxiety in the Holtzman blots (11). Considering the fact that phenomena of anxiety also are reflected in biological functioning a better integration of both the psychological and biological aspects of these bio-social phenomena might become possible by using response variability or consistency as a common index.

SUMMARY

Fifty-four young females were administered three figure drawing tasks

"Personal communication by Dr. Fiske.

twice within an interval of six months. Consistency on a task, the complement of intraindividual variability was measured as related to variation in instructions, time interval between tasks and subjective interpretation of ambiguous instructions. The amount of ambiguity of the stimulus tended to be a determinant of its intensity, clearer instructions constraining variability more than ambiguous ones. Internal psychological stimuli too were found to affect response variability differentially. Results were interpreted as supporting Fiske's notions that intraindividual response variability is constrained by external stimuli, the more so the stronger the stimulus. The principle however seems also to hold for internal psychological stimuli. The possible relationship between intraindividual variability and anxiety as well as the potential usefulness of measures of intraindividual variability in the study of anxiety were discussed.

REFERENCES

- Fiske, D. W. & Rice, L. N. Intra-individual response variability. *Psychol. Bull.* 1955, 52, 217-250.
- Fiske, D. W. The constraints on intra-individual variability in test responses. *Educ. and psychol. Measmt.* 1957, 17, 317-337.
- Fiske, D. W. An intensive study of variability scores. *Educ. psychol. measmt.* 1957, 17, 453-67.
- Fiske, D. W. Variability of responses and the stability of scores and interpretations of projective protocols. *J. proj. Tech.* 1959, 23, 263-267.
- Fiske, D. W. & van Buskirk, C. The stability of interpretations of sentence completion tests. *J. consult. Psychology*, 1959, 23, 177-180.
- Fiske, D. W. & Cox, J. A. Jr. The consistency of ratings by peers. *J. appl. Psychol.* 1960, 44, 11-17.
- Levy, S. Figure drawings as a projective test in Abt, L. E. & Bellak, L. (eds.) *Projective psychology*, N.Y.: Knopf, 1952.
- Lynn, D. B. A note on sex differences in the development of masculine and feminine identification. *Psychol. Rev.*, 1959, 66, 126-135.
- Machover, Karen. Personality projection in the drawing of the human figure, Springfield: C. C. Thomas, 1949.

- Machover, Karen. The drawing of the human figure—a method in personality investigation in Anderson, H. H. & Anderson, Gladys L. (eds.) *Introduction to projective techniques*. N.Y.: Prentice Hall, 1951.
- Mitra, S. K. & Fiske, D. W. Intra-individual variability as related to test score and item. *Educ. psychol. Measmt.*, 1956, 16, 3-12.
- Osterweil, J. & Fiske, D. W. Intra-individual variability in sentence completion responses. *J. abnorm. soc. Psychol.*, 1956, 52, 195-199.
- Shanan, J. Clinical implications of constraints on intra-individual variability in a response to figure drawing tests. Mimeographed. (1960).
- Van Der Veen, F. & Fiske, D. W. Variability among self-ratings in different situations. *Educ. psychol. Measmt.*, 1960, 20, 83-93.

Received May 1, 1961.

The Natural History of an Outcome Prediction^{1, 2, 3}

RICHARD S. SIEGAL, IRWIN C. ROSEN AND GERALD A. EHRENREICH
The Menninger Foundation

From the clinician's point of view the global "will improve vs. won't improve" outcome prediction is less than satisfactory in that it ignores the multidimensional aspects of improvement, bypassing the individual nature of change. We shall report to you a study designed to be clinically natural, to allow the clinician unimpeded use of all the clinical acumen he may possess in predicting psychotherapy outcome for the individual. The Psychotherapy Research Project of The Menninger Foundation is naturalistic in that it attempts to study course and outcome of psychotherapy allowing clinical processes to operate naturally. Neither patient nor therapist, for example, is aware that he is the object of research scrutiny until termination of treatment. Assessments and predictions are made in essentially the same way they would be in service-oriented, daily clinical work. Psychological tests are used as one among a number of ways of collecting data and are used "unnaturally" only in that an attempt is made to be explicit about the bases of inferences and the testing is done "blind" with no extra-test information. The major aim of the project is to determine what has

changed within the individual patient and how the change can be understood in relation to the treatment received. Operating within the psychoanalytic theories of personality and psychotherapy, predictions of the course and outcome of psychotherapy are made and data gathered at the termination of therapy and again two years later to investigate the accuracy of the predictions and the predictive power of the assumptions on which they are based.

The best way to illustrate the research use of clinically natural prediction is to trace the inference chain in an actual case. We are given Rorschach, TAT and Word Association test protocols of a 27-year-old married nurse. By virtue of her inclusion in the research population, we know that she is in some form of psychotherapy, expected on the basis of the clinical work-up to be the major, though not necessarily exclusive, treatment modality. We note all inferences, and the test data upon which they are based, relevant to a list of 22 preselected personality variables, including such things as anxiety and symptoms, core conflict, patterning of defenses, thought organization, affect organization, anxiety tolerance, psychological-mindedness, ego strength and others. Of course, an assessment of one of these variables implies much about assessments of the others. Our assessments of the 22 variables, then, constitute not only a statement about each of them, but a more or less integrated personality description in psychoanalytic terms.

Going to the test protocols we note that the patient, seeing a picture on the office wall, comments that it looks like a European city she once lived in and she tearfully reports this was the city in which she adopted a child she

¹ This paper was presented to a symposium entitled "The Use of Projective Techniques for Predicting the Outcome of Psychotherapy," held at the Annual Meeting of the American Psychological Association, September 1, 1960. The symposium was under the joint sponsorship of Division 12 of the A.P.A. and the Society for Projective Techniques.

² This paper is from the work of the Psychotherapy Research Project of The Menninger Foundation. The generous support of the Foundations' Fund for Research in Psychiatry and the Ford Foundation to this project is gratefully acknowledged.

³ We should like to express our appreciation to Robert S. Wallerstein, M.D., for his critical reading of the paper.

later had to give up because she "couldn't feel like a mother." Later, looking at the TAT picture of the boy with the violin, she again cries, saying, "I can't walk through a baby department in a store or look at baby pictures in a magazine." These and other indications point to the presence of intrusive, depressive concerns restricted, however, to content involving the loss of her adopted child. Nor are the content manifestations of depression accompanied by formal indications of slowed or impaired functioning. Her Word Association responses are given quickly, she recovers rapidly from her tears in the TAT and continues to produce stories, many of which suggest the use of denial in a stout, but losing struggle against depressive affect. There are strong Rorschach indications of efforts to control affect: An Experience Balance of 3 to 5, a Pure Form percent of 77, and a Form Dominant percent of 88. That her efforts to control labile affect fail can be seen in 4 CF responses as opposed to only 2 FC responses. Our inferences from this material emphasize weakening of affect controls leading to labile affect and intrusive depressive ideation. However, depression is not pervasive.

Several test indications point to the prominence, overuse and partial ineffectiveness of ideational activity in the attempt to bind anxiety. Its prominence is suggested by 40 Rorschach responses, including 3 Human Movements. The content of several Rorschach responses (spiders, crabs) suggests the presence of phobic ideas. Her Word Association to *breast* is "child" and she reports a mental picture of the Madonna. This suggested to us the defensive importance of ideational activity particularly in conflictful areas. The erratic functioning of ideation, however, is conveyed in several remarks that her mind "went blank" suggesting she needs to resort to repression when ideation fails sufficiently to bind anxiety. We saw ideational activity, then, as overused and

partially ineffective in binding anxiety. This conveys, along with the ineffectiveness of affect control, a picture of mild decompensation of defenses.

What are the impulses against which her weakening defenses are arrayed? The Word Association *breast*-child with the image of the Madonna leads us to the speculative inference that she needs to idealize and overvalue her mother as a denial of "other," more hostile feelings and attitudes towards her. Throughout the tests the patient manifests a reluctance to associate females with hostile thoughts or feelings. At the same time there are indications of an underlying association between femininity and destructiveness. This is suggested in the content of such Rorschach responses as "a treacherous spider," or a woman's "bony hands." It is strongly confirmed in her TAT stories where while men commit violence, women are interfering, blaming and gossiping. However, each time the theme of the interfering or malicious woman is introduced, the patient quickly disavows it. For example, in one story in which two people are quarrelling about "her mother-in-law who interferes with their lives," the patient says, "I'm just making this up now . . . I have an understanding mother-in-law." The denial and reversal implicit in such comments suggest that this hostile conception of women (ultimately mother) is especially threatening. Specifically, we infer from this and similar responses that she experiences mother unconsciously not only as one who interferes, blames and gossips, but as dangerous and malevolent. She intensely though unconsciously resents mother yet feels that she is like mother, destructive and evil. She must avoid becoming conscious of her own hostile and resentful ideas about mother as well as her identification with her.

Her characterological ways of dealing with the anxiety stemming from these unconscious conflicts can be seen in test indications of previously

effective reaction formation and counterphobic activities. Her heightened anxiety to gain control of her mother suggests the presence of reaction formation. A reaction like *dirty mother* is never purely its characterological availability. Also characterologically, she can find acceptance of her mother's effect in active, counterphobic roles. Her response to the mother's threat is "insect powder." She comments, "dirty mother, I hate bugs. I do first thing get out my insect powder and get rid of them—I hate bugs." Both reactions to the mother's threat and her counterphobic tendency can find expression in the assumed support of professional role of the nurse. Thus, while she is seriously threatened in her decompensation with the possibility of unconscious being, she may feel her mother becoming conscious she has a characterological reversal. This suggests the possibility that reversal of the decompensation with a return to established character patterns may be a plausible reaction to her acute conflicts.

In summary then, we infer that ideational activity fails to bind anxiety effectively enough, affect controls are weakened and extremely disturbing, intrusive depressive concerns are present. We see this as the result of partial decompensation of the defense mechanism of reaction formation and the ineffectiveness of denial and reversal called upon to support reaction formation. We see the patient threatened by the possibility of regressive impulses entering consciousness, impulses conveying her perception of women, her mother and herself, as evil and destructive. Her failure as a mother, we feel, is unconsciously perceived as a punishment of that part of her which is identified with her own mother. We feel, however, that reversal of the decompensation, re-establishment of reaction formation and the possibility of her reassuming her counterphobic role may obviate the necessity, for her, of becoming conscious of these repressed and feared

hostile impulses.

In essence we have described certain aspects of the patient's behavior under more or less known conditions, the tests. Our problem is to extrapolate from this conceptualization to a set of different circumstances, psychotherapy. From this point on our inferences are somewhat removed from the test data and their usefulness or validity seems likely to rest squarely upon the psychologist's experience with and understanding of psychotherapy. Our job at this point in the prediction process is to decide, on the basis of our personality assessments, what kind of psychotherapeutic intervention is appropriate for psychopathology of this nature and extent. The question is one of predicting the differential effects of a given psychotherapy from within the range of psychoanalytically-oriented psychotherapeutic modalities.

While it is not possible here to discuss the indications for different therapeutic modalities, the questions involved are two. The first of these is: What psychotherapeutic modality is necessary to ameliorate the disturbance? We inferred decompensation of defenses affecting ideational functioning and control of affect, in a woman unable to function effectively and comfortably in a mature feminine role. For such pervasive psychopathology a major, intensive psychotherapeutic endeavor is necessary. We felt therefore that psychoanalysis would be necessary for a stable resolution of conflict. The second question is: Given the particular state of affairs in the patient and the known or hypothesized processes of particular psychotherapies, what psychotherapy will be possible? What problems and limitations may be foreseen in a given modality? Gill (2) implies this double question in his statement, "The choice of therapy may be divided into that which determines the minimum necessary to restore the ego-functioning, and that which strives for the maximum change that is possible." A major

aspect of psychoanalysis which differentiates it from other psychotherapeutic approaches is a pull toward more regressed functioning. As Malcalpine (3) puts it, regression is encouraged by the very structure of the psychoanalytic situation which brings about the development of a regressive transference neurosis repeating the original infantile situation out of which later pathology developed.

How will these aspects of psychoanalysis impinge upon this patient's character structure and psychopathology? The answer to this is, in essence, our outcome prediction. Outcome predictions, we believe, are simply arbitrarily isolated, sometimes disguised, predictions of therapeutic process frozen in time. It is a truism in psychotherapy research that outcome and process are interlocked and inseparable, yet researchers often try to predict therapy outcomes without consideration of therapeutic processes.

This woman's defensive efforts are directed toward keeping unconscious her feelings, rooted in infantile perceptions, that mother was malicious and destructive, that she is like mother, and that this is what it means to be a woman. The anxiety stemming from the emergence of these ideas into consciousness is intense. The "pull" to regress in psychoanalysis will again threaten the effectiveness of her defenses and the resulting anxiety will be particularly intense. While such anxiety associated with the threatened emergence of unconscious ideas is, of course, usual, our generic outcome prediction is that at a certain point in psychoanalysis this anxiety will become so intense and influential she will feel continued exploration of unconscious conflict too threatening. This will impose significant limits upon the analytic achievements she will realize.

This general prediction is amenable to extension and translation into more specific forms. For example, a limitation in insight is predictable. Specifically, we think it likely the pa-

tient will achieve only limited, perhaps intermittent, and fragmentary awareness of her unconscious perceptions of mother as malicious, vengeful and destructive. The limitation will be apparent in another way. Ordinarily the transference would be the primary vehicle for expressing these unconscious conflicts. We believe the patient will continue to feel too threatened by these deepest conflicts in relation to mother to allow the unconscious attitudes to emerge in the transference situation. The likelihood, therefore, for the expression of intense negative transference reactions is small. While resistance against the emergence of these unconscious ideas could take manifold forms, most consistent with our understanding of her character structure would be the re-institution of reaction formation and the possible resumption of her counterphobic professional role. At the point at which she begins to feel unduly threatened she may quit while she is ahead, and settle for the gains she has already realized. This does not imply a precipitous or very early termination, nor that her achievements will not be significant, but from the point of view of the usual aim of psychoanalysis—reasonably complete resolution of major conflicts, it might be considered premature.

The many knotty and unsolved problems in confirming or refuting such predictions cannot be dwelt upon here. We think we are aware of them. Many predictions go wrong. Some, hopefully, are confirmed. With sufficient explanation of the various steps in the prediction process there is always the possibility, even if the prediction is incorrect, of identifying the locus of error. Simply to demonstrate that such predictions are in principle confirmable and that their specificity is appropriate, and to encourage at least a minimal feeling of closure, we present the following material collected from termination interviews with the patient, the analyst and the supervisor of the case, by an

independent team of researchers with no knowledge of our predictions. Of the transference the analyst said "She verbalized almost exclusively the fact that the nature of her relationship to me was that of considering me the father. There were many evidences that this was a front behind which was the yearning to make a good, kind, protective and loving mother." The research group considering the nature of the termination said: "The patient had also come to a point in treatment where the therapist was seen as the protecting, kindly mother and the breaking up at that time could be seen as a way to keep him in that perspective." The therapist said of the outcome of treatment: "There were significant changes but not as far as my expectation." The patient said: "I have not gained in three years what I hoped I would gain, but I am still more comfortable." The supervisor commented: "I would describe it as a clinical success with moderate to marked improvement. However, I believe the transference was never handled fully."

To summarize, we have attempted to explicate a process of building inferences stemming originally from psychological tests, into a chain, joining our assessments of intrapsychic variables with inferences based on hypotheses about psychotherapeutic processes, finally to arrive at a generic outcome prediction. From this we

derived further predictions relevant to certain major aspects of treatment.

We have tried to make each step on our predictive path visible, especially in the crucial area where our understanding of intrapsychic phenomena and psychotherapeutic processes converge. We feel that psychotherapy research which short-circuits this explication or settles for unnecessarily global outcome predictions deprives the clinical researcher of the chance to use and enhance his clinical acumen and predictive sophistication.

REFERENCES

- Ehrenreich, Gerald A. and Siegal, Richard S. Assessing Intrapsychic Variables with Psychological Tests in Psychotherapy Research. To be published.
- Gill, Merton. Ego Psychology and Psychotherapy. *Psa. Quart.*, 1951, 20, 60-71.
- Macalpine, Ida. The Development of the Transference. *Psa. Quart.*, 1950, 19, 501-539.
- Robbins, Lewis L. and Wallerstein, Robert S. The Research Strategy and Tactics of the Psychotherapy Research Project of The Menninger Foundation and the Problem of Controls. In Eli A. Rubinstein and Morris B. Parloff (Eds.) *Research in Psychotherapy*. Washington, D.C. American Psychological Association, 1959, Pp. 27-43.
- The Psychotherapy Research Project of The Menninger Foundation: Rationale, Method and Sample Use: First Report. *Bull. Menninger Clinic*, 1956, 20, 221-278.
- The Psychotherapy Research Project of The Menninger Foundation: Second Report. *Bull. Menninger Clinic*, 1958, 22, 115-166.
- The Psychotherapy Research Project of The Menninger Foundation: Third Report. *Bull. Menninger Clinic*, 1960, 24, 157-216.

Received August 11, 1961.

Manifestations of Primary Process Thinking in Schizophrenia¹

LLOYD H. SILVERMAN

Department of Psychiatry, University of Pennsylvania and New York University

BENJAMIN LAPKIN

Hillside Hospital

AND IRA S. ROSENBAUM

College of the City of New York

In studying schizophrenia, particularly in recent years, psychoanalytic writers have focused on ego impairment. Considering the ego as a substructure of personality defined by its functions, Beres (1956) and Bellak (1958) have enumerated those they considered to be severely impaired in this disorder: (1) relation to reality (2) regulation and control of instinctual drives (3) object relations (4) thought processes (5) defensive functions (6) the autonomous functions and (7) the synthetic function.

In discussing the disturbance in *thought processes* in schizophrenia, Bellak writes "The formal characteristics of schizophrenic thinking are identical with the formal descriptive characteristics of the primary process." (1958, p. 19) Holt has described these characteristics in the following way: "(Primary process thinking) has certain peculiar formal characteristics. These include autistic logic instead of straight thinking, loose and nonsensical types of associative links and distortion of reality in numerous ways. But the most notable formal deviations in primary thinking were described by Freud as the mechanisms of the dream work. *Condensation* is a process resulting in a fusion of two or more ideas or images. *Displacement* is a shift of emphasis or interest from one mental content to another (usually a less important content in terms of relevance to conflict or instinctual

aims). *Symbolization* is the replacement of one idea or image by another, always a concrete visual presentation which may have various formal features in common with what is being symbolized, but which disguises the latter's dynamic significance" (1956, pp. 15-16).

The Rorschach has proven to be one of the most effective instruments for detecting and studying the formal aspects of primary process thinking, particularly starting with Rapaport's most important contribution to the literature in 1946. He examined and discussed Rorschach manifestations of this kind of thinking through a detailed and systematic examination of the *verbalization* of responses. Rapaport stated, "The S's verbal communication of his response in relationship to the determinants he used, the area he chose, and the content he attributed to it, is an integral aspect of that response: like the response itself, it is a product of the subject's thought processes and is amenable to scoring, systematization, and diagnostic evaluation." (1946, pp. 324-325). Rapaport then listed and illustrated various kinds of "deviant verbalizations" or what we would call responses in which the formal aspects of primary process thinking are noteworthy. Rapaport considered such responses to be particularly prevalent in the Rorschachs of schizophrenics, and backed up this claim both with rational considerations and with some supporting data.

In at least three subsequent studies, these deviant verbalization categories were tested and were found to be highly differentiating. Watkins and

¹ We wish to express our thanks to Dr. Jacob Cohen for his advice on the statistical procedures employed and to Mrs. Rita A. Simon and Dr. Irving Steingart for their assistance in other phases of the study.

Stauffacher (1952) assigned weights to the different kinds of deviant verbalizations which Rapaport listed, and developed a "delta index" as a measure of pathological thinking. They found that this index clearly differentiated a group of 25 psychotics, both from a group of 25 normals and from a group of 25 neurotics. In 1955 Powells and Hamlin tested the differentiating power of this index on 10 normals, 10 anxiety neurotics, 10 latent schizophrenics, 10 paranoid schizophrenics, and 10 catatonic schizophrenics. All three schizophrenic groups had significantly higher index scores than the normal and neurotic groups. More recently, Hertz and Paulino (1960) compared 35 male paranoid schizophrenics with 35 male neurotics of approximately the same age and IQ scores. In their comparisons, the point of departure was the kind of organizational patterns that characterized each group, rather than the thought processes per se. Their general hypothesis was confirmed, namely that the paranoid schizophrenics to a greater degree than the neurotics, "showed organizational patterns reflecting ego impairment and disturbance in experiencing the ambiguous perceptual material of the Rorschach" (1960, p. 374). Since some of the specific organizational patterns investigated were operationally defined in terms of the occurrence of the kinds of indicators of pathological thinking that were first introduced by Rapaport (particularly the "fabulized combination" score), these results too, are relevant.

Holt, in the early 1950s, started the development of a manual to assess the formal and content characteristics of primary process thinking which he has reported on in a series of papers (1954, 1956, 1960a). His discussion of formal characteristics has much in common with Rapaport's listing and, in fact, certain of Rapaport's categories were adopted by Holt. The manual seems to us to be an advance over both Rapaport's

listing and the Watkins and Stauffacher delta index for the following reasons: (1) It specifies more of the formal characteristics of primary process thinking since it was specifically developed to assess such thinking. (2) The different categories of primary process formal characteristics are defined in a more detailed way than heretofore has been the case and are well illustrated with many examples. (3) Two provisions are made for assessing the degree to which the formal aspects of primary process thinking are reflected in a response. First, scored responses are designated as Level 1 or Level 2, depending on whether the degree is judged to be large or small respectively. Secondly, a "defense demand" (DD) rating is given for each scored response which more finely assesses the degree of primary process influence. A six point scale is used with the higher numbers reflecting greater primary process influence, or in Holt's words, "(the greater) the shock value of the response . . . (the more it) demands that some defensive and controlling measures be undertaken in order to make it a socially acceptable communication" (1960 b, p. 62). (4) There is also provision for evaluating the degree to which the S is master of the primary process elements in his thinking in a "defense effectiveness" (DE) rating. This has been defined as the rating given to a scorable primary process manifestation for "the effectiveness of controlling and defensive measures in reducing or preventing anxiety and making a successful adaptive response to the examiner's demand to interpret the blots" (1960 b, p. 63). DE, like DD ranges along a six point scale, the higher the number assigned, the less effective these measures.

Accounting for degree of defense effectiveness seems most desirable in light of the growing acceptance of the concept of regression in the service of the ego. Holt writes, "It is one of mankind's great gifts to be able to aban-

don reality voluntarily for a little while; to shake free from dead literalism, to re-combine the old familiar elements into new imaginative amusing or beautiful patterns . . . (Kris) has pointed out the fact that the ego of a mature and healthy person can at times relax, abandon secondary process standards in a controlled and recoverable way and use the freedom and fluidity of the primary process productively . . . The person who is not asleep and dreaming may therefore, fragment and re-combine ideas and images in ways that flout the demands of reality on either of two bases; because he cannot help it due to a temporary or permanent weakness, or because he wants to for fun or for creative purposes and is able to because he is not too threatened by his unconscious drives." (1956, p. 16). Holt's DE score is intended to help differentiate these two kinds of individuals.

Holt's manual has been used to assess primary process thinking in a number of studies. Goldberger (1958) used it to predict individual differences in reaction to perceptual isolation. Pine and Holt (1960) studied the relationships between expression and control of primary process material, on the one hand, and quality of productions in a variety of tests of imagination on the other. Lapkin (1960) used the manual to study which qualities of thinking are related to the recovery of subliminal material. Both Philip (1960) and Saretsky (1961) made use of it to study the effects of drugs on thinking. Von Holt et al (1960) related manual scores to performance on the Hanfmann-Kasarin test. And finally Ackman (1960) used the manual for comparing thinking under normal conditions with thinking when Ss are under hypnosis and when they are given "fantastic" instructions.

Up to this point, no attempt has been made to use the manual for comparing different diagnostic groups for primary process thinking. In light of

some of the distinguishing characteristics of schizophrenia as discussed earlier, such a comparison seems very much in order. It was with this in mind that our study was undertaken.

METHOD

Subjects

Ss were 40 adolescents (ages 11 to 18) who were patients at the Hawthorne Cedar Knolls and Linden Hill schools, open residential treatment centers that are divisions of the Jewish Board of Guardians. They were divided into two groups, one consisting of 20 schizophrenics and the second of 20 psychoneurotics and personality disorders (as these terms are used in the diagnostic manual of the American Psychiatric Association (1952)). All 40 Ss were patients for whom there was concurrence on diagnosis by at least two clinicians, usually the psychiatric social worker who was treating the patient and a psychiatrist². The Ss were matched in pairs for I.Q., age, and sex. There was no significant difference between the mean ages of the schizophrenic and neurotic-personality disorder groups (15.7 and 15.8 respectively) or between the mean I.Qs. (105.8 and 103.4). There were 17 males and three females in each group.

Procedure

The Rorschach was administered individually to each S, according to the procedure described by Rapaport (1946). The protocols were coded and were assessed "blindly" by two of the writers with each of them evaluating an equal number of patients in each group. The records were scored for primary process thinking according to

² The psychiatrists and social workers, in most instances, probably had read psychological reports on the patients they had diagnosed sometime in the past. In order to minimize the possibility of "contamination," at least one of the two clinicians assessing each patient was specifically asked to base his diagnosis *only* on material that he *directly* obtained in sessions with the patient.

the degree of control in the Holt manual. The manual also allows for the possibility of a patient's not being able to differentiate between the two groups of patients. This is a possibility which is not taken into account in the manual. The manual also allows for the possibility of a patient's not being able to differentiate between the two groups of patients. This is a possibility which is not taken into account in the manual. The manual also allows for the possibility of a patient's not being able to differentiate between the two groups of patients. This is a possibility which is not taken into account in the manual.

Hypothesis I

Hypothesis I is based on the expectation that schizophrenics will manifest in their thinking more of the formal characteristics of the primary process than will normals and personality disordered. Since the Holt manual had not been used before to differentiate schizophrenics from other kinds of patients we did not know whether to anticipate differences both for Level 1 and Level 2 formal aspects or just for Level 1 (the more extreme). Thus hypothesis I is broken down into the following two sub-hypotheses: (a) The schizophrenic records will be given more Level 1 and Level 2 scores for formal aspects. (b) The schizophrenic records will be given more Level 1 scores for formal aspects.

Hypothesis II is also concerned with the formal aspects of primary process thinking but in light of the earlier discussion of regression in the service of the ego, it takes into consideration additionally the Ss degree of mastery over these manifestations. Here too, two sub-hypotheses are formulated: (a) The schizophrenic records will be given more Level 1 and Level 2 scores for formal aspects that are given a DE

score of three or more. (b) The schizophrenic records will be given more Level 1 scores for formal aspects that are given a DE score of three or more.

Hypothesis III When the records of the two groups of patients are examined globally, primarily based on a factor, involvement of the variables does not differ in hypotheses I and II. This can be distinguished from each other on a better than chance basis.

Hypothesis IV In addition to providing for the assessment of the formal characteristics of primary process thinking the Holt manual also allows for an evaluation of content characteristics that is, the degree to which the content of thought bears the stamp of a libidinal or aggressive drive. Schizophrenics have been described as having poor control over their instinctual drives (Beles, 1956; Bellak, 1958) and we thought that such poor control might show itself in drive domination as reflected in either of two ways, (1) the libidinal or aggressive idea would be extreme and in manual terms would be scored for Level 1 rather than Level 2 content, (2) the libidinal or aggressive idea would be given in a relatively maladaptive fashion and in manual terms would be given a DE rating of three or higher. Hypothesis IV thus consists of the following sub-hypotheses: (a) The schizophrenic records will be given more Level 1 libidinal scores, aggressive scores, and total content scores (i.e., libidinal and aggressive scores combined). (b) The schizophrenic records will be given more scores with a DE of 3-6 for Level 1 libidinal content, aggressive content, and total content. (c) The schizophrenic records will be given more scores with a DE of 3-6 for Level 1 and Level 2 libidinal content, aggressive content, and total content.

Since it has been hypothesized that when the formal aspects and the content aspects of primary process thinking are considered separately, the schizophrenics will be given more

form F forms and more content with a DE. All this is to be expected, then, in the response and content of schizophrenics and normal individuals.

Subsidiary Hypothesis 1. Although not a major concern for this study, is the comparison schizophrenics with normal and personality disorders for various ratings of primary process thinking. The design also allows us to test the secondary hypothesis. One of these forms at poor form level which is usually by writers have described as characteristic of schizophrenia. Rorschach (Koster, 1942; Schacter, 1948; and Holt, 1960, 1964). The Holt manual includes criteria for scoring form level and this is one response characteristic, but has to be assessed in deriving a DE score. The scores used are those proposed by Maymin (1960) in his Rorschach Form Level Manual. The scale: F^{+} , sharp, convincing forms; F , popular and near popular forms; Fw , reasonably plausible but not terribly convincing forms.

F^{-} , forms that bear only a slight resemblance to the blot area; F^{-} , arbitrary forms; Fv , vague non-definite forms; Fa , amorphous responses in which form plays no role and Fs , spoiled form responses.

Holt (1960, p. 73). In our first subsidiary hypothesis, we make the following two predictions: (a) The schizophrenics will give more responses where the form level is F^{+} , Fs , or Fa ; (b) Using the weightings proposed by Holt (1960, b, p. 73) with F^{+} getting the highest weight and F^{-} the lowest, the schizophrenics will have the lower average weight.

Our second subsidiary hypothesis is concerned with those schizophrenics who give Rorschach records with relatively few responses reflecting the formal and content aspects of primary process thinking over which the S is not the master. We suspect that these patients can manage this only by defensively clinging to reality bound and drive-distant secondary process thinking. We expect this clinging to show

up in both an overabundance of responses of going for primary process thinking and in some of the presence of some of the following indicators: (1) Many responses to the response to reality; (2) Many F responses; (3) Many F forms; (4) Fw , Fv , Fs , Fa responses; (5) Many formal, nonprimary responses. Holt's (1960, p. 73) and (8) Rorschach manual includes criteria for scoring content and this is one response characteristic, but has to be assessed in deriving a DE score. One of the indicators by which we will test our secondary hypothesis will be that schizophrenics will make a mistake in the primary process hypothesis. With all 40 Ss asked to make up the degree to which they may have been wrong for formal content with DE at 40 and wrong for content with DE at 40. In the Ss in the bottom half of the distribution who are schizophrenics will have higher "clinging to reality" scores than the normal and personality disorders in the bottom half of the distribution.

(b) Those Ss in the bottom half of each distribution who are schizophrenics will have lower average DD scores for whichever primary process aspect is involved in a formal or content than the normal and personality disorders in the bottom half of the distribution.

RESULTS

The mean number of Rorschach responses is 20.0 for the schizophrenics and 16.7 for the normal and personality disorders (with standard deviations of 10.30 and 5.51 respectively). A t test for the difference between the two means produced a value of 1.57 which is not significant. However, since the schizophrenics gave more re-

3 This indicator was scored for such remarks as "this is an inkblot" or "this is a line down the middle" and for the S focusing on some concrete object in the testing situation, for example, asking the examiner "Are you using a Schaeffer pen?"

TABLE I—Formal Aspects of Primary Process Thinking

	Schizophrenics		Neur. & Pers. Dis.		t value	p
	Mean	S. D.	Mean	S. D.		
Level 1 and Level 2 Density Scores						
(Equal Weighting)	71.15	50.20	43.40	27.57	2.34	.02
(Differential Weighting)	102.60*	72.33	64.50	42.77	1.98	.03
Level 1 Density Scores						
(Equal Weighting)	18.80	24.09	10.80	14.52	1.24	—
(Differential Weighting)	27.00	28.31	16.15	16.45	1.44	.09

* The mean % of R can be greater than 100 since any response can be given more than one score.

TABLE II—Formal Aspects of Primary Process Thinking Not Well Mastered

	Schizophrenics		Neur. & Pers. Dis.		t value	p
	Mean	S. D.	Mean	S. D.		
Level 1 and Level 2 Density Scores						
DE 3-6						
(Equal Weighting)	48.60	49.01	22.45	23.52	2.09	.025
(Differential Weighting)	70.95	74.75	32.95	38.60	1.97	.03
Level 1 Density Scores						
DE 3-6						
(Equal Weighting)	16.50	24.11	7.40	12.29	1.46	.08
(Differential Weighting)	21.25	26.01	9.65	14.70	1.69	.05

sponses, we felt that the more stringent test of our hypotheses would involve comparisons in which the total number of Rorschach responses for each S was taken into consideration. The unit that we chose to employ for making these comparisons reflects the "density" of primary process⁴.

The results bearing on Hypothesis I (the density scores for formal aspects of primary process thinking) are reported in Table I⁵. When both Level

1 and Level 2 formal aspects are considered, the two t values are significant⁶. When only Level 1 formal aspects are considered, the t values, although in the predicted direction, are not significant. However, the t value, when weak scores are weighted $\frac{1}{2}$, indicates a tendency in the predicted direction (i.e., it would be significant if the criterion were relaxed to the .10 level).

The results bearing on Hypothesis II (the density scores for formal aspects that are not well mastered) are reported in Table II. When both Level

⁴ This "density score" was obtained by dividing R into the number of scores given for primary process thinking. It is a different unit than that which usually has been used in studies employing the Holt manual in which R has been divided into the number of responses with one or more scores for primary process thinking. We chose the density score as the unit to work with since it reflects the varying number of primary process scores that different responses can be given. When we also compared the two groups for the other, more usually used unit, the differences were smaller in most instances than when the density scores were compared. The fact that the latter better discriminated between the schizophrenics and non-schizophrenics in this study argues for its use in other studies where the Holt manual is employed.

⁵ The Holt manual distinguishes between "full" scores and "weak" scores. Since the manual has not been used to compare diagnostic groups before, in this table and in

Tables II, IV, and V, we experimented in our treatment of these scores, and calculated totals in two ways: Where both Level 1 and Level 2 responses are involved, the first mean density score is based only on Level 1 full, Level 1 weak, and Level 2 full scores, all of these being given equal weight. The second mean density score is based on weighting Level 1 full scores double, Level 1 weak and Level 2 full scores singly, and Level 2 weak scores $\frac{1}{2}$. Where only Level 1 responses are involved, the first mean density score is based on only full scores. The second mean density score is based both on full scores and weak scores, the latter being given a weight of $\frac{1}{2}$.

⁶ For this table and those that follow, results are presented in terms of a one tail test.

TABLE III—Diagnostic Judgments

Actual Status	Schiz.	Judgment of Status				Total Neur. & Pers. Dis.
		Don't Know Schiz.	Total Schiz.	Neur. & Pers. Dis.	Don't Know Neur. & Pers. Dis.	
Schizophrenics (N=20)	8	5	13	4	3	7
Neurotics and Personality Disorders (N=20)	5	2	7	8	5	13

TABLE IV—Drive Dominated Content Scores

	Schizophrenics		Neur. & Pers. Dis.		t value	p
	Mean	S. D.	Mean	S. D.		
Level 1 and Level 2						
Lib. Density Scores DE 3-6						
(Equal Weighting)	5.15	6.96	7.50	9.34	.88	—
(Differential Weighting)	8.05	9.53	9.35	10.45	.40	—
Level 1 and Level 2						
Aggr. Density Scores DE 3-6						
(Equal Weighting)	15.10	11.53	11.55	11.17	.96	—
(Differential Weighting)	19.30	15.27	14.05	12.98	1.14	—
Level 1 and Level 2						
Tot. Content Density Scores DE 3-6						
(Equal Weighting)	20.45	16.32	19.50	17.26	.17	—
(Differential Weighting)	26.90	22.34	24.05	21.19	.40	—
	Number of Schiz. Ss with One or More Scores*		Number of Neur. & P. D. Ss with One or More Scores			
Level 1 Lib.	5		1			—
Level 1 Aggr.	3		3			—
Level 1 Tot. Content	5		3			—
Level 1 Lib. DE 3-6	4		1			—
Level 1 Aggr. DE 3-6	1		3			—
Level 1 Tot. Content DE 3-6	4		3			—

* Since most Ss in both groups were not given any scores for the six latter categories, we felt there was no point in comparing mean density scores. The chi-square contingency test was employed to test for the differences between the number of Ss in the two groups who were given one or more scores for each of the categories.

1 and Level 2 scores are considered, the two t values are significant. When only Level 1 scores are considered, the t value is significant when weak scores are counted, but there is only a tendency in the predicted direction when only full scores are counted.

The results bearing on Hypothesis III (the global judgments of whether or not a patient is schizophrenic), are presented in Table III. When all the judgments are considered, 26 of them are correct and 14 incorrect. These results are significant (binomial test). When the "don't know" but "forced choice" judgments are excluded, 24 cases remain in which the evaluators

felt reasonably sure of their predictions. They were correct in 16 of these judgments and incorrect in 8. These results are not significant⁷.

⁷ It is somewhat surprising that the global judgments are significantly differentiating when all 40 cases are considered but not so when only those cases are considered where the evaluators were reasonably certain of their judgments. While the per cent of correct judgments in both instances is approximately the same, indicating that it is only the reduction of N in the latter instance that precluded significant results, we expected this reduction to be more than compensated for by the elimination of the two "don't know" categories. However, some of the edge is taken off this unexpected finding when it is noted that one of the evaluators, although being quite unsure about four of

TABLE V—Total Primary Process Scores

		Schizophrenics		Neur. & Pers. Dis.		t value	p
		Mean	S. D.	Mean	S. D.		
Level 1 and Level 2 Density Scores							
DE 3-6							
(Equal Weighting)		68.65	61.14	41.95	39.98	1.59	.07
(Differential Weighting)		97.05	90.90	56.45	56.41	1.65	.06
Level 1 Density Scores							
DE 3-6							
(Equal Weighting)		17.85	24.15	8.65	13.77	1.44	.09
(Differential Weighting)		22.60	26.03	11.40	16.58	1.58	.08
Level 1 Density Scores							
(Equal Weighting)		21.20	24.46	12.05	15.72	1.37	.09
(Differential Weighting)		29.60	28.75	19.25	18.55	1.32	.10

TABLE VI—Form Level Scores

	Schizophrenics		Neur. & Pers. Dis.		t value	p
	Mean	S. D.	Mean	S. D.		
% of Responses						
Scored F—, Fs, or Fa	7.55	8.08	9.35	10.78	.58	—
Sum of Form						
Level Weights	40.97	55.37	42.45	54.43	.83	—

TABLE VII — "Clinging to Reality" Scores and DD Scores for Low Primary Process Subjects

	Mean "Clinging to Reality" Score	Mean DD
Schizophrenics with Low Formal Scores with DE 3-6 (N=8)	80.40	1.04
Neur. & Pers. Dis. with Low Formal Scores with DE 3-6 (N=12)	77.70	.87
Schizophrenics with Low Content Scores with DE 3-6 (N=11)	76.00	.88
Neur. & Pers. Dis. with Low Content Scores with DE 3-6 (N=10)	80.40	.81

The results bearing on Hypothesis IV (the density scores for drive dominated content) are presented in Table

IV. None of the t values is significant.

The scores for total primary process are presented in Table V. While none of the t values is significant, they all indicate a tendency in the predicted direction. This is a natural consequence of the fact that each total score is made up of a formal score, for which there is a significant difference between the groups, and a content score, for which there is not.

The results bearing on the first of the subsidiary hypotheses (form level) are presented in Table VI. Neither of the t values is significant. The results bearing on the second subsidiary hypothesis (clinging to secondary process thinking) are presented in Table VII. Here, too, none of the t values is significant.

DISCUSSION

Our hypotheses bearing on the ability of the Rorschach to detect the impairment in thought processes of schizophrenics were largely supported by the findings. Thus, the schizophrenic records had both significantly greater density scores for all Level 1 and Level 2 formal manifestations of primary process (Hypothesis I) and for those formal manifestations that were not well mastered (Hypothesis

his cases, finally included them in the unequivocal categories since he was *somewhat less* unsure about them than he was about cases he placed in the "don't know" categories. Actually, he felt that they should be categorized somewhere between these two kinds of categories. If these four cases had been included among those eliminated, 21 cases would have remained of which 15 were judged correctly, a finding that is significant.

II)^{8,9}. In addition, when the Rorschach protocols were evaluated globally for diagnostic purposes, with a noteworthy degree of disturbance in the thought processes considered as the major indicator of schizophrenia, the two groups could be differentiated at least when all 40 records were under consideration. These findings are very much in keeping with earlier studies of Rorschach manifestations of

⁸ The fact that these hypotheses were less strongly supported when only Level 1 formal aspects were considered we believe to be largely due to the fact that the schizophrenics as well as the neurotics and personality disorders gave few responses that received such scores which, in turn, meant that differences between the two groups had to be quite large in order to be significant. The average number of Level 1 formal scores for the schizophrenics was 5.8 compared to 13.7 Level 2 formal scores. Similarly, the average number of such Level 1 scores with DE of 3-6 was 4.1 compared to 9.3 Level 2 formal scores with DE of 3-6. One possible explanation for the absence of a greater number of Level 1 scores in our sample may be related to the fact that these schizophrenics were able to be maintained in a residential treatment center at the time they were treated. Perhaps Level 1 formal manifestations are only found in large numbers in hospitalized schizophrenics. A second possibility is that it is the age of the schizophrenic that is crucial. Perhaps the greater degree of pathology reflected in the Level 1 formal aspects only comes about with increasing chronicity and is not found to any great extent in adolescent schizophrenics. A study is being planned in which hospitalized adult schizophrenics will be evaluated in terms of the Holt manual, our expectation being that they will be given significantly more Level 1 scores for formal aspects of primary process thinking than the schizophrenics in the current study.

⁹ In light of the earlier discussion of regression in the service of the ego, it is noteworthy that there was a significant difference between the two groups for formal manifestations regardless of their DE rating. We suspect that this was because the group with which the schizophrenics were compared did not contain many, if any, Ss who were capable of adaptive regression in their thinking. If a schizophrenic group were compared with a group of creative Ss, for example, it would be our expectation that the difference between them (or at least the greater difference between them) would be in terms of formal scores with DE of 3-6 rather than formal scores per se.

thought disturbance in schizophrenic adults (Rapaport, 1946; Watkins and Stauffacher, 1952; Powers and Hamlin, 1955; and Hertz and Paolino, 1960) and indicate that such disturbance is also detectable in the Rorschachs of schizophrenic adolescents. The findings of significant differences in this investigation seems to us to be of particular note in light of the particular characteristics of the two groups studied. The schizophrenic Ss did not require hospitalization and thus do not represent the "sickest" segment of the schizophrenic population, while the neurotics and personality disorders *did* require residential treatment, suggesting that they have greater ego disturbance than most patients with similar diagnoses. Thus, one would have expected a minimizing of the differences in ego functioning that usually exist between schizophrenics on the one hand and neurotics and personality disorders on the other. The fact that there were still significant differences between the groups therefore attests to the differentiating power of the scores for formal aspects in the Holt manual.

The hypothesis bearing on control over instinctual drives was unsupported by our findings. However, a post hoc reevaluation of the data indicated that for control over aggressive drives, (though not for control over libidinal drives), some support for the hypothesis was in evidence if the data were examined in ways other than that which was originally planned. Thus, if the groups are compared for the percent of responses with any Level 1 or Level 2 aggressive scores (full or weak), with a DE of 3-6, there is a tendency in the predicted direction ($p = .08$). And if the *number* of these responses (rather than the percent) is examined, the difference between the groups is significant¹⁰.

Since these findings were arrived at

¹⁰ It was mentioned earlier that since the schizophrenics gave more (though not significantly more) responses, the more con-

post hoc (and thus are in need of cross validation), they offer only weak support for the hypothesis that schizophrenics manifest poor control over their aggressive drives. However, they closely parallel the results of an earlier investigation of the sorting test performance of the same 40 subjects that were studied here (Silverman and Silverman, in press). In that study the schizophrenics, as was hypothesized beforehand, gave significantly more responses reflecting intrusions of aggressive drive¹¹. The findings on aggression from both studies seem to us to be particularly notable in light of the fact that by comparing the schizophrenics with the particular group of neurotics and personality disorders that we used, we lessened our chances of supporting this hypothesis. For, of the 20 neurotics and personality disorders, 18 were placed in residential treatment, either wholly or partly because of various kinds of overt aggressive behavior¹². Thus, there is independent evidence that these Ss may have been impaired in their ability to control aggressive drives to a greater extent than would be expected, both in non-psychiatric groups and in neurotic-personality disorder groups

servative test of the hypotheses would take R into consideration. However, the number of scores per se also can be of value. In evaluating Rorschach records for clinical purposes, there are instances when particular manifestations take on crucial diagnostic importance regardless of the number of responses given.

¹¹ Actually, the results, as they were reported in this paper, were in terms of the schizophrenics giving more responses reflecting intrusions of aggressive and libidinal drives taken together. However, a further breakdown of this finding subsequently undertaken indicates that this overall difference is clearly due to the difference in aggressive intrusions. The fact that in both studies it was poor control over aggressive rather than libidinal drives that was suggested will be discussed in a future paper.

¹² In the intake summaries, assaulting others was mentioned for nine of the patients, stealing for nine, defiant behavior for six, temper tantrums for six, destroying property for three, and a suicide attempt for one.

which did not require residential treatment. We would anticipate that comparing schizophrenics with either of these latter groups would provide clear-cut support for our hypothesis regarding control over aggressive drive.

Neither of our subsidiary hypotheses was supported. In regard to poor form level, the absence of significant differences may have been a function of the particular sample of schizophrenic and neurotic-personality disorder patients we used, which as was described above, probably led to a reduction of the usual difference in ego functioning between these diagnostic types. At any rate, subject to confirmation by other investigators, it would appear that Rorschach responses reflecting the formal aspects of primary process are more sensitive indicators of schizophrenia than responses with poor form.

The negative findings concerning the ability of certain schizophrenics to avoid primary process thinking by clinging to secondary process thinking can be ascribed to either (or both) of two possibilities. The formulation itself can be erroneous or we could have chosen wrong or insufficient Rorschach variables to focus on. In a future study, we hope to return to this issue.

SUMMARY

A comparison was made of the Rorschachs of a group of 20 schizophrenic adolescents with 20 psychoneurotic and personality disorders, matched for age, I.Q., and sex. The major hypotheses called for a "blind" evaluation of test responses for formal and content aspects of primary process thinking. The results supported the expectations that the Rorschachs of the schizophrenics; (1) would contain significantly more manifestations of thinking that bear the formal characteristics of the primary process and (2) contain significantly more poorly mastered manifestations of this kind of thinking. However, the expectation that the schizophrenic Rorschachs

would be more drive dominated received only weak support, and then only for aggressive drive and only when the data were reworked post hoc. The findings were discussed, relating them to other studies of Rorschach manifestation of primary process thinking in schizophrenia.

REFERENCES

- Ackman, Phyllis. The effects of induced regression on thinking processes. *Dis. Abs.* 1960, 21, 365.
- Ainsworth, Mary. Problems of validation. In: Klopfer, Ainsworth, Mary; Klopfer, B.; & Holt, R. R. *Developments in the Rorschach Technique*. New York: World Book Co., 1954.
- American Psychiatric Association. *Diagnostic and statistical manual, mental disorders*. Washington: Mental Hospital Service, 1952.
- Bellak, L. (Ed.) *Schizophrenia: a review of the syndrome*. Chap 1. New York: Logos Press, 1958.
- Beres, D. Ego deviation and the concept of schizophrenia. *The Psychoanalytic study of the child*, 1956, 11, 164-235.
- Goldberger, L. Individual differences in effects of perceptual isolation as related to Rorschach manifestations of the primary process. Unpublished doctoral dissertation, New York University, 1958.
- Hertz, Marguerite & Paolino, A. F. Rorschach indices of perceptual and conceptual disorganization. *J. proj. Tech.*, 1960, 24, 370-388.
- Holt, R. R. *The influence of primary and secondary process on Rorschach content*. Paper presented at N.Y. Chapter of Soc. Proj. Tech., 1954.
- Holt, R. R. Gauging primary and secondary processes in Rorschach responses. *J. proj. Tech.*, 1956, 20, 14-25.
- Holt, R. R. & Havel, Joan. A Method for assessing primary and secondary process in the Rorschach. In Maria A. Rickers-Ovsiankina (Ed.), *Rorschach psychology*. New York: Wiley, 1960. (a)
- Holt, R. R. *Manual for the scoring of primary process manifestations in Rorschach responses*. Draft 8. New York University Research Center for Mental Health, 1960. (Dittoed) (b)
- Klopfer, B. & Kelley, D. M. *The Rorschach Technique*. New York: World Book, 1942.
- Lapkin, B. The relation of primary process thinking to the recovery of subliminal material. Unpublished doctoral dissertation, New York University, 1960.
- Mayman, M. *Rorschach form level manual*. Menninger Foundation, 1960. (Dittoed)
- Phillips, A. E. The effect of lysergic acid diethylamide (LSD-25) on primary process thought manifestations. Unpublished doctoral dissertation. New York University, 1960.
- Pine, F. & Holt, R. R. Creativity and primary process: a study of adaptive regression. *J. abnorm. soc. Psychol.*, 1960, 61, 370-379.
- Powers, W. T. & Hamlin, R. M. Relationship between diagnostic category and deviant verbalizations on the Rorschach. *J. consult. Psychol.*, 1955, 14, 120-124.
- Rapaport, D. *Diagnostic psychological testing*. Vol. 2, Chicago: Year Book Publishers, 1946.
- Saretsky, T. The effect of chlorpromazine on primary process thought manifestations. Unpublished doctoral dissertation, New York University, 1961.
- Schafer, R. *Clinical application of psychological tests*. New York: International Universities Press, 1948.
- Silverman, L. H. & Silverman, Doris K. Ego impairment in schizophrenia as reflected in the object sorting test. *J. abnorm. soc. Psychol.*, in press.
- Von Holt, H. W., Sengstake, C. B., Sanoda, B. C., & Draper, W. A. Orality, image fusion, and concept-formation. *J. proj. Tech.*, 1960, 24, 194-198.
- Watkins, J. & Staffacher, J. An index of pathological thinking. *J. proj. Tech.*, 1952, 16, 276-286.

Received August 7, 1961

BOOK REVIEWS

Brodin, Henry W. (Ed.) *Lectures on Experimental Psychiatry*. Pittsburgh: University of Pittsburgh Press, 1961. Pp. 361, \$7.50

The present volume is a series of lectures, most of them given at a conference on experimental psychiatry at the Western Psychiatric Institute in Pittsburgh in 1959.

The format of lectures is not, however, the best for a book. The format of a lecture is to deliver a certain amount of material, to leave the audience to read and think about it. The format of a book is to deliver a certain amount of material, to leave the reader to read and think about it. The format of a lecture is to deliver a certain amount of material, to leave the audience to read and think about it. The format of a book is to deliver a certain amount of material, to leave the reader to read and think about it.

The same rule, unfortunately, holds true in the field of psychiatry. From reading the present book, one gets the impression that the material was delivered at the conference, and that the reader is to read and think about it. The format of a lecture is to deliver a certain amount of material, to leave the audience to read and think about it. The format of a book is to deliver a certain amount of material, to leave the reader to read and think about it.

One of the apparent problems in putting together a book such as this is to delimit the bounds of the subject matter, i.e., what constitutes experimental psychiatry. The chapters and subject matter included in this book range from paralinguistic to brain surgery to lecture trips abroad. Hence, one is left with an amorphous impression and an uneasy feeling about the nature and the limit of the subject matter. Perhaps this is a reflection of the state of affairs in the field of psychiatry, perhaps the field of psychiatry is trying to include too much as has been pointed out by writers such as Szasz. But all this should have been discussed and made clear in such a series of lectures.

This book is of little interest to psychologists for three reasons. First, the book is too heavily loaded with neurophysiological, physiological and drug research. Approximately one-third of the book deals with such topics. However, it is not implied here that such pieces of research, in themselves, are not worthwhile contributions. Second, most of the researches which are reported in the book essentially deal with a particular phase

of a research project or simply give a few of the preliminary results of a piece of research. Such partial reporting of research doesn't give a feeling of closure and leaves the reader with unfulfilled expectations. Third, there is really very little in the way of experimentation and research in this book which is not already known to those who keep up with the literature. Most of the chapters on research and experimentation are naive and global, and simply discuss a few of the basic principles of setting up experimental designs.

Nevertheless, there are some rewarding chapters in the book. Benjamin's article, "Innate and the Experiential in Child Development" discusses beautifully the interaction between these two aspects of development and clearly shows how certain maturational (Innate) crises, can be either reduced or exacerbated by the experiences that one encounters. Liddell's article on "Conditioning in the Sheep and Goat" shows that the process of conditioning is an emotional experience for the organism and "it involves primarily the animal's struggle to fathom the meaning of what is happening and about to happen to him" (pp. 235). Gerard's article, "Knowledge and Neural Functions," is a thought-provoking article dealing with the growth and development of knowledge in any field of science.

It is my understanding that a book is usually published for at least one of the four following reasons: first, to disseminate new knowledge or to present a viewpoint; second, to consolidate, at one place, a large amount of scattered knowledge or studies which are not easily accessible and which seem to be the order of the day; third, to systematize and categorize knowledge; and finally, to stimulate thinking and curiosity amongst the readers. For this reviewer, it is difficult to determine for which of these purposes this book was published.

SOHAN LAL SHARMA, PH.D.
Senior Staff Psychologist,
Los Angeles Psychiatric Service

Eysenck, H. J. (Ed.). *Handbook of Abnormal Psychology. An Experimental Approach*. New York. Basic Books, 1961. Pp. XVI + 816. \$18.00.

This is one handbook which appears to merit the name of handbook. It is a weighty

and comprehensive coverage of experimental work in the field of abnormal psychology. The pages are double column pages with abundant illustrations, graphs and tables. While the illustrations and graphs are printed throughout most of the volume, it is at the same time a scholarly and critical work. Most of the authors are English psychologists, who have, or have had, some affiliation with the Department of Psychology at the University of London. Some are now functioning in other countries. Cyril M. Franks of the New Jersey Neuropsychiatric Institute, appears to be the only one who is currently working in the United States. In addition, a few chapters have been authored by colleagues in psychiatry.

The book consists of 20 chapters divided into three main parts. There is also a brief foreword by William A. Hunt and an introduction by the editor. The three main sections of the book are: Description and Measurement of Abnormal Behavior (Part I), Cause and Determinants of Abnormal Behavior (Part II), and Experimental Study and Modification of Abnormal Behavior (Part III). It may be worthwhile to list the separate chapters as indicators of the content of the book, since one cannot really describe or evaluate the contents of this volume adequately in a brief review. Part I consists of seven chapters: Classification and the Problem of Diagnosis (H. J. Eysenck), Abnormalities of Psychomotor Functions (A. Yates), Expressive Movements and Abnormal Behavior (H. Brengelmann), Abnormalities of Sensory Perception (G. W. Ganger), Intellectual Abilities and Problem-solving Behavior (W. D. Furneaux), Cognitive Abnormalities (R. W. Payne), and Abnormalities of Motivation and "Ego-functions" (J. Inglis). The second part includes: Heredity and Psychological Abnormality (J. Shields and E. Slater), Constitutional Factors and Abnormal Behavior (L. Rees), Childhood Upbringing and Other Environmental Factors (N. O'Connor and Cyril M. Franks), Somatic Reactivity (I. Martin), Conditioning and Abnormal Behavior (Cyril M. Franks), and Learning and Abnormal Behavior (H. Gwynne Jones).

The third part, which has more direct reference to clinical functions, consists of the following chapters: Psychological Effects of Brain Damage (V. Meyer), The Effects of Psychosurgical Procedures on Behavior (R. Willett), The Psychological Effects of Cerebral Electroshock (D. Campbell), The Effects of Drugs on Behavior (D. Trouton and H. J. Eysenck), The Effects of Psychotherapy (H. J. Eysenck), Abnormal Animal Behavior (P.

Broadhurst), and Applied Abnormal Psychology: the Experimental Approach (H. Gwynne Jones).

As indicated in the subtitle of the book, this is a volume on experimental studies in abnormal psychology. As can be noted, from the chapter headings, also, a wide array of areas of abnormal behavior which have been studied experimentally are included in the presentation. As one might anticipate from a volume organized and fathered by H. J. Eysenck, there is also a certain critical as well as crusading or polemical tone about the volume. Eysenck, in particular is quite critical of purely clinical or subjective approaches to the study of abnormal behavior. He is particularly critical of the use of projective techniques, and of the use of the term "psychiatric diagnosis" and the use of such terms as "abnormal behavior." He is quite sure that his views on this matter and the position of psychologists will be sympathetic to his point of view in this regard—although they may not be inclined to go all the way with him, or to accept his solution to the problem. As mentioned in his introduction, he and his colleagues "have rejected the psychiatric framework outright, preferring not to be bound and constrained by its concepts, notions of dubious validity and unreliability." (xv). Although the authors are not using the psychiatric terminology, they have tried to translate these terms into the experimental and views suggested by Eysenck. As long as clinical and experimental psychology are content to take psychiatric diagnosis as their criterion, so long will their experimental designs be of exemplary simplicity and elegance, and the results equivocal and uninterpretable. (xv). It should be added, however, that in spite of such disclaimers, there is still a fair amount of the volume which summarizes research organized around conventional psychiatric categories. There is at the same time considerable concern in many of the chapters with evaluation of studies in terms of adequacy of the experimental design and the lack of cross-validation.

In terms of this frame of reference, Eysenck states that most of the published material had to be discarded since it was not useful. Much of the literature utilizing projective techniques apparently falls into this category. "Less than one fifth of the papers and books examined have been included, and even so the reader will find that what remains often gives rise to critical discussion rather than to enlightenment. We have preferred the positive to the negative approach, leaving out

whole areas which traditionally would have been included (the vast literature of the Rorschach is one example) because detailed investigation revealed all dross and no gold, and including instead much less well cultivated areas (such as dark adaptation, or conditioning) because these seemed to hold out a more definite promise of future usefulness" (xi-xii).

Eysenck strongly emphasizes that his conception of abnormality is in terms of the defective functioning of psychological systems, rather than "in terms of classes of people suffering from mental diseases produced by different 'causes'" (xii). To the present reviewer the point of view expressed by Eysenck is clearly that of an experimental psychologist rather than that of the clinician who is interested in the individual *qua* individual. Although my own professional orientation is inclined more toward the latter, I see much justification for Eysenck's point of view and for his critical attitude toward much of clinical practice and research. The reliance on psychiatric diagnosis in much of our research has certainly been a limitation leading to a wide variety of conflicting results, as is amply demonstrated throughout this volume.

The general attempt throughout the book is "to see abnormal psychology as part of general, experimental psychology. The laws of learning theory, to take one example, apply no less to neurotics than to rats and college students; should it not be the task of the psychologist to deduce the details of abnormal behavior from such general laws" (xiii). There is little question that Eysenck and his collaborators have attempted to adhere to this orientation in their presentations.

I have tried thus far to present the general framework and guiding philosophy of the volume. Because of space, it will not be possible to comment adequately on each of the chapters individually. I also do not feel competent to evaluate critically all of the areas covered. Instead, I will comment primarily on a few sections or chapters which appear to hold greatest interest for most clinical psychologists and then to give some overall evaluation of this work.

The first chapter on "Classification and the Problem of Diagnosis" should be of interest to clinicians although they will not all agree with Eysenck's point of view. In general, Eysenck attacks the prevailing psychiatric system of nosology and does a competent job in terms of his criticism. The lack of knowledge concerning cause of psychiatric disturbances, problems in the use of the disease

entity concept, lack of reliability, and similar criticisms are presented in a telling manner.

Eysenck offers as a substitute for reliance on psychiatric classification, the analysis of the behavior of patients, utilizing factor analysis. In terms of this procedure, he has developed his system of "dimensional analysis". The distinguishing features of this approach to diagnostic study lie in the use of quantifiable objective methods of appraisal, the factor analysis of specific test data, and the securing of factors or dimensions which can be utilized to describe different types of patient behavior. Eysenck has come up with three main dimensions along which patients can be ordered. One of these is introversion-extraversion; another is neuroticism and a third is psychoticism. When psychiatrically diagnosed patients are ordered along these dimensions, one finds that there is continuity along these dimensions and also that many of the conventional diagnostic categories can be grouped in a meaningful manner. For example, groups of normals, neurotics, and psychotics can be differentiated in terms of the neuroticism and psychoticism factors. Furthermore, when a two dimensional framework is used utilizing the dimensions of neuroticism and introversion-extraversion, not only are normals and neurotics differentiated but individual groups of neurotic patients are also distinguished. The introverted neurotic grouping has been labeled *dysthymia* by Eysenck and the patients diagnosed as anxiety states or obsessive neurosis fall within this group. Individuals who are high in neuroticism and also high on extraversion tend to be diagnosed as hysterics or psychopaths. Patients who are high on the neuroticism scale but intermediate on the introversion-extraversion dimension tend to be diagnosed as "mixed neurosis". Eysenck claims several advantages for this approach. "Instead of a variable number of disease entities, differing from psychiatrist to psychiatrist, and diagnosed only with low reliability, we have two dimensions along which every person can be ranged and given a numerical score. Continuity is thus substituted for discontinuity and measurement for discrete classification" (p. 10). In addition, it would appear that testable hypotheses and interrelationships are obtainable from this orientation without undue difficulty.

Most of the other chapters tend to have the same orientation although they vary somewhat in this regard. One chapter which appears to come closest to the area of projective techniques is the chapter on expressive movements. In addition to a critical discussion

of such bodily expressions as gait, gesture, posture, facial expression, voice, etc., the author also reviews the literature on drawing, handwriting and related techniques. Much of the latter material has some relevance for projective theory and includes also research findings on such tests as the Draw-A-Person and the House-Tree-Person Tests. The author's review of the findings on such techniques is concise, critical and thought provoking. He points out that although one may agree that free drawing and other artistic products may reveal aspects of personality, this view has no scientific merit "until we have demonstrated, *firstly*, that expressive movement *per se* is diagnostic of personality, and not via intelligence, artistic excellence, age, etc. *Secondly*, the basic measures of expressive movement, their functional conditions and personality correlates have to be defined in order to map out the field. The inescapable conclusion, after perusal of the literature, is that no one of these basic aims has even been approached during decades of analysis by hundreds of investigators" (p. 73). In addition to an evaluation of free drawing techniques there is also an evaluation of "semi-objective techniques", such as the Bender-Gestalt and the Lowenfeld Mosaic Tests, and some comparable, but not as well known, types of tests developed by European psychologists. The latter include the Colour-Pyramid Test, the Colour-Star Test and the Bunch-of-Grapes Test. Included also in this chapter is a test by the author of the chapter (H. Brengelmann), called the Figure Reconstruction Test. This test measures learning, recall and recognition as well as expressive movement and requires the reproduction of exposed patterns of simple geometric shapes (circle, triangle, square, etc.) to be drawn from memory. Performance on this test is objectively scored and has been related to measures of extraversion, rigidity, hysteria-dysthymia, schizophrenia and a measure of confidence in performance. In general, most of the tests reviewed in this chapter are found wanting as currently used. According to the author the main factors accounting for this are poor (clinical) validating criteria, heterogeneous test scores based on logical rather than on empirical grounds, and a disregard of the impact of conditional effects on test performance.

I would also like to comment specifically on a few other chapters of interest before making some general concluding statements about the volume as a whole. The chapter on "Cognitive Abnormalities" by Payne surveys a considerable amount of literature pertain-

ing to intellectual performance and impairment in varied clinical groups. Much of this relates to the intellectual level and degree of impairment found in several types of psychiatric patients. Although the author points out that (at least for psychoses) "psychiatric diagnoses do not have the status of scientific concepts", he is forced to present data based on such concepts. Numerous studies are then reviewed based on tests of intelligence and on special tests of impairment. Most of the conclusions reached are not astounding, e.g. neurotic disorders produce very little deterioration, but psychosis produces a considerable degree of deterioration. However, his conclusion that "schizophrenics are duller than normal before becoming ill" (p. 209) may be more controversial. These conclusions are based on general test results from different studies and at times it is not always easy to follow the author's deductions. In other instances the author has lumped together scores from different studies of schizophrenics and attempted to draw conclusions from such data. As most of the authors are aware, unreliability of diagnosis, duration and extent of disturbance, age, and other variables make generalizations somewhat hazardous. The second half of the chapter is more original and thought provoking. Concepts such as variability of performance, test content, type of problem (e.g. perceptual, inductive reasoning, deductive reasoning, fluency, originality), and cognitive processes are utilized to analyze a large amount of test results. Some of the results of this analysis do not support some conventional clinical views concerning intellectual patterns in psychotic patients. The section on cognitive processes follows in great part the rationale provided by Furneaux in the previous chapter. Here such concepts as speed of response, persistence, error, distractibility, memory span, learning, retention, rigidity, and concept formation become the points of reference for analyzing intellectual performance. This type of analysis appears to offer something to our way of analyzing and understanding the cognitive processes of psychotic patients. It would appear, for example, that lowered performance on some tasks can be explained more parsimoniously by the time limit of the test than by the name given to it by the test author.

Somewhat briefer comments can be given for a few other chapters of particular interest to workers in this field. The chapter on "Heredity and Psychological Abnormality" appears to be a good summary of material in this area—an area which clinical psychologists at times are apt to minimize or disregard.

The chapter concerning childhood experience is probably of greater interest. This rather brief chapter is a critical evaluation of current concepts, views and beliefs in this area, with some emphasis on aspects of psychoanalytic theory. The lacks and deficiencies in this area are clearly pointed out. The chapter by Eysenck on "The Effects of Psychotherapy" is quite in line with what clinicians would have anticipated from his 1952 critique of psychotherapy in the *Journal of Consulting Psychology*. Although this area represents a special interest of the reviewer, the purpose of the review does not permit any lengthy discussion. It may be of interest to point out that Eysenck is very critical of the research reported by Rogers and Dymond, in their *Psychotherapy and Personality Change*. Actually, Eysenck is even more impressed with the negative results derived from psychotherapy than he was earlier. However, he has a more positive view toward "behavior therapy", advanced by such workers as Wolpe, and which is based on learning theory. In terms of this, he appears much less critical of the learning theory utilized than is true of some American psychologists, e.g. Rotter and Mowrer. Finally, the last chapter of the book "Applied Abnormal Psychology: The Experimental Approach" by H. Gwynne Jones, is of some interest. Jones makes a plea for the psychologist to justify his professional existence by making an independent contribution to the clinical field. From his view "the special contribution of psychologists should be the application of the findings and techniques of general psychology to the problems involved in the investigation and treatment of psychiatric patients" (p. 764). While he emphasizes objective and experimental approaches to patients, he does present an interesting and appealing rationale. Essentially, this involves studying the patient with specific techniques selected for the particular problem and then utilizing or devising appropriate therapeutic or remedial techniques for helping the patient. While one may not agree completely with the techniques mentioned by Jones, this approach has a strong logical appeal. Apart from the matter of objectivity, our current diagnostic and therapeutic techniques appear to be used as if they have almost universal application to all kinds of diverse problems.

As has been indicated, this is a book which is concerned with the experimental analysis of various types of abnormal behavior. The topics listed and discussed previously give some indication of the content. While some of the material would appear to consist of

basic knowledge for many persons working in the field of abnormal behavior, it is clearly not a book devised to enable the psychologist to better understand the actual problems of a disturbed individual. Most of the text deals with experimental studies of abnormal groups of subjects and the emphasis is thus on group findings and on objective methods of appraising behavior. A moderate amount of the work is also organized around the dimensional schema of Eysenck. The book thus differs from *Personality and the Behavior Disorders* edited by J. McV. Hunt in 1944, to which some may compare it. Although the latter work covers somewhat the same field, the two volumes differ in terms of how they handle and select the data. One can say that the Hunt volume had more material of a clinical nature and had a greater diversity of viewpoint. The current volume is much more typical of the older texts in abnormal psychology in which the chapters were arranged in terms of disorders of perception, disorders in thinking, etc. The volume also has a particular viewpoint which is highly critical of current approaches and points of view in clinical psychology and psychiatry. This message is hammered out repeatedly in almost all of the chapters. While some may feel that this criticality is almost obsessively overdone, one should not disregard all of the criticisms because of this. Unfortunately, Eysenck's provocative style may turn away some readers who might profit from contact with the book.

SOL L. GARFIELD
Nebraska Psychiatric Institute
University of Nebraska

Huber, Jack T., "Report Writing in Psychology and Psychiatry", Harper & Brothers, 1961, X+114, \$3.50

This book presents elements of psychological report writing, and a few examples of psychiatric report writing in a terse, compact, and quite elementary manner. Many nuggets of wisdom are presented in the remarkable variety of aspects which are covered in such a small volume. These are co-mingled, however, with over-simplifications such as the reassurance, that, after all, each psychologist (and psychiatrist) will develop his own style of reporting. A number of pages are devoted to citing report writing outlines proposed by various writers including the author. Indeed, one of the chief merits of the book is the compilation of material from a variety of sources which, although for the most part

readily obtainable in the original publications, are conveniently brought together here. Huber's own contribution is a practical, chronological-topical presentation outline form which gives due regard to developmental factors. To then conclude that the student is his own best arbiter of report forms hardly follows. Similarly, the student is told he need not worry about theory—everyone, the author states, has his own; nor need he concern himself overly much with suggesting a diagnosis—many clinics, the student is informed, specifically avoid diagnosing because of the antiquated psychiatric nomenclature. Elsewhere, he writes that a professional opinion is an opinion that is given with assurance. These over-statements are modified after their initial bald statement. The reviewer cites them as example of dicta proposed to beginning report writers which might better be dealt with in a text on psychodiagnostics or on clinical psychology.

The book invites comparison with another on the same topic, published in 1960, "The Psychological Report" by Walter Klopfer, Grune & Stratton, (erroneously titled in Huber's bibliography). Klopfer presents a method of teaching the student to organize a report by deriving the data in a systematic distillation from the test responses. Huber starts with the organizing of the report. Klopfer gives the reader a feeling for the depth of psychodiagnostics and psychological report writing in clinical psychology. Huber gives preference to the style of report writing suited to industrial psychology. He commits a measure of *lese majesté* in critically analyzing case report excerpts from Beck, Schafer, and Bellak because of their technical psycho-analytic language, taking these excerpts out of their context in advanced-student text books in which setting they were highly communicative. Both books are handy manuals and worthwhile additions to the clinician's shelf, the one for depth, and the other for wide-range breadth of content.

LEONARD S. ABRAMSON
VA Hospital, East Orange, N. J.

Kataguchi, Yasufumi, (Ed.) *Rorschachiana Japonica*, Vol. IV, 1961. Tokyo Institute of Rorschach Research, Tokyo, Japan.

The fourth consecutive annual of *Rorschachiana Japonica* has been published and, like the previous three, under the able editorship of Yasufumi Kataguchi. The yearbook contains sixteen (16) original papers on the Rorschach written by Japanese psychologists, preceded by a Japanese translation of Pauline Vorhaus' biographical sketch of Bruno Klopfer, who is a consulting editor to the yearbook.

A number of papers deal with Rorschach pictures of such clinical groups as delinquents, schizophrenics, psychosomatics, and epilepsy. There are studies of psychotherapy, children, solitary confinement, the Rosenzweig, P-F test, form-level rating, blind analysis, prognostic rating scales and ego psychology.

The scope of the studies is broad and the references, including American ones, more up-to-date than is customary in foreign publications. While it is apparent that some Japanese psychologists lack modern library facilities, most of them are actively engaged in the same research front that we are.

The utility of this series is somewhat limited for other than the Japanese students (for whom it is intended) because it is printed in Japanese in contrast with some other Japanese journals. About one-half of the articles have English summaries. Of some interest is the fact that many tables and graphic presentations are in English and therefore permit English speaking psychologists to view and evaluate the statistical data. I believe that the editors would be doing a service were they in future editions to provide English summaries of all the articles. They would find many interested readers in this country.

BERTRAM R. FORER
Los Angeles

ANNOUNCEMENTS

WORKSHOP

1962 WORKSHOP IN THE RORSCHACH TECHNIQUE OF PERSONALITY DIAG- NOSIS AND OTHER PROJECTIVE TECH- NIQUES AS USED WITH CHILDREN

jointly sponsored by

Claremont Graduate School in Claremont
and Children's Hospital, Los Angeles
directed by

BRUNO KLOPFER and
HELMUT WURSTEN

Time and Place: July 8 to 20, 1962 at
Ashomar Conference Grounds, Pacific Grove,
California

The Workshop will be devoted to the study
of projective techniques as used with chil-
dren. It is planned for professional people
who are competent in the basic use of the
Rorschach Test. An opportunity will be given
to examine and discuss the test protocols of
children presenting a wide variety of dis-
orders.

Courses: (The number of two-hour sessions
is indicated after each course):

1. Introductory and Survey Lectures (4)
2. Review of Administration and
Scoring (10)
3. Case Study Seminar, including Case
History Analysis, Psychometric Material
and Other Projective Material (20)

Claremont Graduate Credit: Qualified
graduate students accepted for the Workshop
may register for two units of credit (Psy-
chology 243, or 244a,b,c.)

All students wishing to qualify for gradu-
ate credit must apply to the Claremont Grad-
uate School, Claremont, California, for the
necessary forms before May 15, 1962, after
they have been notified of admission to the
Workshop.

Tuition: \$60.00

Applications: Write to Dr. Bruno Klopfer,
P.O. Box 2971, Carmel, California, before
June 1, 1962.

TWO WORKSHOP SEMINARS IN THE RORSCHACH TEST

Department of Psychology
The University of Chicago
Summer 1962

- The Foundations:* Administration. Proces-
sing the associations. Psychologic signifi-
cance of the test variables. The whole
personality. Introduction to interpreta-
tion. June 18-22, 1962.
- Advanced clinical interpretation:* Adults
in acute stress. The younger child in
neurotic conflict. The test in mental defi-

ciency. Symptoms, intellectual, in schizo-
phrenia. Treatment problems in an ado-
lescent disorder. June 25-29, 1962.

Dr. S. J. Beck will conduct both semin-
ars. For information write to: Rorschach
Workshops, Department of Psychology,
University of Chicago, Chicago 37, Illinois

REPRINTS AVAILABLE

The paper by G. Blum: *A guide for re-
search use of the Black Pictures*, is available
as a separate reprint with cover for \$1.25
from the Psychological Corporation, 304 East
45th St., New York 17, N.Y.

Because of the sparsity of detailed clinical
cases for training and education in the use
of projective methods, the Society for Projective
Techniques prepared a symposium:
the *Case of El* which was published in three
parts and presented in part at the 1961 con-
vention of the American Psychological Asso-
ciation. The entire case study has been assem-
bled and will be available as a training
monograph at \$2.50. Contents are as follows:

Part 1. Edwin S. Shneidman: Test protocols

1. Rorschach Psychodiagnostic Technique
2. Murray Thematic Apperception Test
3. Shneidman Make-A-Picture-Story (MAPS)
Test
4. Forer Sentence Completion Test
5. Shneidman Interest-Completion Test
6. Strong Vocational Interest Blank
7. Kuder Vocational Preference Record
8. Guilford-Shneidman-Zimmerman
Interest Survey
9. Allport-Vernon-Lindzey Study of Values
10. Hildreth Feeling and Attitude Scale
11. Cornell Selectee Index
12. Hathaway-McKinley Minnesota Multi-
phasic Personality Inventory
13. Shipley-Hartford Retreat Scale
14. Wechsler Adult Intelligence Scale
15. Watson-Glaser Test of Critical Thinking
16. Tsedek Test of Moral Judgment

Part 2: Biography of El written by Evelyn
Hooker, the therapist.

Part 3: Test interpretations.

Forer, Bertram R. Vocational Choice
Fine, Reuben. The MAPS Test and TAT.
Meyer, Mortimer M. A blind analysis
Rorschach and TAT.

Shneidman, Edwin S. The logic of El: A
psycho-logical approach
Murray, Henry A. Commentary on the
case of El.

To order your copy of the *Case of El*, send
\$2.50 to the Society for Projective Techniques,
210 E. Wilson Ave., Glendale 6, Calif.

DIRECTORY OF MEMBERS OF THE SOCIETY FOR PROJECTIVE TECHNIQUES AS OF MARCH 1, 1962

Date preceded by A indicates date elected as Associate
Date preceded by AM indicates date elected as Associate Member
Date preceded by AM indicates date elected as Associate Member
Date preceded by F indicates date elected as Fellow
Date preceded by H.F. indicates date elected as Honorary Fellow
Name preceded by * indicates Charter Member or Fellow

- ABEL, Theodora Mead (Ph.D.)**
Palisades A 1944
Rockland County, N.Y. F 1945
- ABRAMS, Elias N. (Ph.D.)**
178 E. 59th St.
Brooklyn 3, New York A 1912
- ABRAMS, Jules C. (Ph.D.)**
1505 Paper Mill Road
Philadelphia 18, Pa. A 1955
- ABRAMS, Julian (Dr.)**
Psychology Department
Springfield State Hospital
Sikesville, Md. A 1954
- ABRAMS, Ray H. (Ph.D.)**
408 S. Lansdowne Avenue
Lansdowne, Pa. A 1950
- ABRAMSON, Leonard S. (Ph.D.)**
Psychology Serv. V.A.H. A 1948
East Orange, N.J. F 1955
- ABT, Lawrence Edwin (Ph.D.)**
151 Rockland Avenue
Larchmont, N.Y. A 1951
- ACKERMAN, Bernard R.**
124 F. 91st Street
New York 28, N.Y. A 1943
- ADAMS, Henry B. (Ph.D.)**
McGuire VA Hospital
Richmond, Va. A 1957
- AINSWORTH, Mary D. (Ph.D.)**
Dept. of Psych.
Johns Hopkins Univ. A 1948
Baltimore 18, Md. F 1950
- AKAVIA, Mr. Uriel**
P.O. Box 11193
Tel Aviv, Israel A 1959
- AIRCOCK, Miss Theodora**
3 Devonshire A 1950
London, W1, England F 1953
- ALEXANDER, Herbert M.**
547 Barbara Drive
Norristown, Pa. A 1958
- ALEXANDER, William A.**
Dept. of Psychology
Western Reserve Univ.
Cleveland 6, Ohio A 1955
- ALIEN, Doris Twitchell (Ph.D.)**
2447 Cloybourne Pl.
Cincinnati 19, Ohio A 1949
- ALIEN, Robert M. (Ph.D.)**
Department of Psychology
University of Miami A 1949
Miami 46, Florida F 1951
- *ALOZERY, Jessie Jervis (Ph.D.)**
314 E. 41st St.
New York 17, N.Y. F 1940
- ALTMAN, Charlotte Hall**
Institute for Juvenile (Ph.D.)
Research
907 South Wolcott Avenue
Chicago 12, Ill. F 1956
- AMCHIN, Abraham**
8 Lyons Ave.
So. Farmingdale, N.Y. A 1952
- AMES, Mrs. Louise Bates**
Gezell Institute of Child
Development
510 Prospect Street
New Haven 11, Conn. A 1951
- ANDERSON, Helen Joan**
Suite 429, Hotel Bonairt
94 Montague Street
Brooklyn 1, N.Y. A 1949
- APPELL, Melville J.**
388 Browncroft Blvd.
Rochester 9, New York A 1950
- ARMON, Mrs. Mary Virginia**
470 W. Avenue 43
Los Angeles 65, Calif. A 1946
- ARNAUD, Sara H. (Ph.D.)**
Arsenal Family &
Children's Center
40th St. and Penn Ave.
Pittsburgh 24, Pa. A 1958
- ARONSON, Marvin L. (Ph.D.)**
225 E. 73rd Street
New York 21, N.Y. A 1951
- ASHER, Miss Miriam**
241 West 4th St.
New York 14, N.Y. A 1960
- ATCHISON, Calvin Oglin**
Box 442, Tennessee (Ed.D.)
St. Univ.,
Nashville, Tenn. A 1961
- AUERBACH, Mrs. Aline B.**
440 East 56th Street
New York 22, N.Y. A 1944
- AZIMA, Mrs. Fern Cramer**
Allan Institute
1025 Pine Avenue W.
Montreal, P.Q., Canada A 1951
- BACHRACH, Arthur J. (Ph.D.)**
Div. of Behavioral Science
University of Virginia
School of Medicine A 1950
Charlottesville, Va. F 1954
- BAKER, Corinne F. (Ph.D.)**
1839 Farmington Road
East Cleveland 12, O. A 1943
- BAKER, Gertrude (Ph.D.)**
2726 Montana Avenue
Santa Monica, Calif. A 1948
- BAKER, Lawrence M. (Ph.D.)**
Department of Psychology
Purdue University
Lafayette, Indiana F 1956
- BALKEN, Eva Ruth (Ph.D.)**
133 East 58th Street
New York 22, N.Y. A 1948
- BARAHAL, George D. (Ph.D.)**
Clinical and Ed. Psychology
Wayne University
Detroit 1, Mich. A 1954
- BARBARA, Dr. Peter Paul**
Box 31
Hall Brooke Hospital
Green Farms, Conn. A 1952
- BARCLAY, Allen (Ph.D.)**
6185 Pershing Ave.
St. Louis, Mo. A 1960
- BARKLEY, Bill J. (Ph.D.)**
1654 Coventry Road
Cleveland, Ohio A 1950
- BARON, Louis K.**
230 N. Tenth Street
Philadelphia 41, Pa. A 1951
- BARON, Samuel (Ph.D.)**
310 West 86th St.
New York 24, N.Y. A 1944
- BARRFELI, Robert P. (Ph.D.)**
Veterans Department
Veterans Admin. Hospital
Downey, Ill. A 1952
- BARRINGER, Benton E. (Ph.D.)**
219 Pinecrest Dr.
Fairview, No. Car. A 1951
- BARRY, John R. (Ph.D.)**
College of Health Related
University of Florida
Gainesville 16, Fla. A 1953
- BARTLETT, Mrs. Doris A.**
924 West End Avenue
New York 25, N.Y. A 1949
- BAUGHMAN, Emmett Earl**
Dept. of Psychology (Ph.D.)
Univ. of North Carolina
Chapel Hill, N.C. F 1957
- BEALE, Elizabeth A., Ph.D.**
4609 Russell Ave.
Los Angeles 27, Calif. A 1951
- BEARDSLEY, Katherine (Ph.D.)**
Clinical Psychology Branch
St. Elizabeth's Hosp.
Washington 20, D.C. A 1950
- BEATTY, Miss Eleanor**
Child Study Inst.
Bryn Mawr College
Bryn Mawr, Pa. A 1958
- BEAUCHEMIN, Jean M.**
10795 Esplanade
Montreal 12, Quebec
Canada A 1952
- BECK, Samuel J. (Ph.D.)**
Department of Psychology
University of Chicago
Chicago 37, Ill. F 1950
- BEISEL, Mrs. Lillie Burling**
104 Beach Avenue
Larchmont, N.Y. A 1944
- BELL, John E. (Ed.D.)**
1112 Hillview Road
Park Hills
Berkeley 8, Calif. A 1949
F 1951
- BELLAK, Leopold (M.D.)**
1160 Fifth Avenue
New York 29, N.Y. A 1948
F 1952
- BENE, Eva (Ph.D.)**
1 Upper Wimpole St.
London, W1, England F 1961
- BERAN, Marianne (Ph.D.)**
408 First Avenue
New York 10, N.Y. A 1949

- BERGSTROM-BORLAND, Mrs. Ingrid**
1025 Las Orejas Ave.
San Rafael, Calif. A 1956
- BERLINER, Anna (Ph.D.)**
Pacific University
Forest Grove, Ore. A 1954
- BERLINER, Mrs. Hildegard**
120 Commonwealth Ave.
San Francisco 18, Cal. M 1959
- BERNSTEIN, Mrs. Hilde R.**
5470 Hydepark Blvd.
Chicago 15, Ill. A 1953
- BERNSTEIN, Lewis (Ph.D.)**
3720 East 7th Ave.
Denver 6, Colo. F 1961
- BERNSTEIN, Louis (Ph.D.)**
Robt. Bruce Apts., Annex
B-10, 300 Hatboro Pike
Hatboro, Pa. A 1957
- BERNSTEIN, Mildred R. (Ph.D.)**
498 Hempstead Ave.
Malverne, L.I., N.Y. A 1950
- BERRICK, Myron E. (Ph.D.)**
1086 Ocean Avenue
Brooklyn 30, N.Y. A 1955
- BIELIAUSKAS, Vytautas J.**
Xavier University
Cincinnati, Ohio F 1956
- BILLIG, Otto (M.D.)**
Department of Psychiatry
Vanderbilt University
Hospital A 1941
Nashville, Tenn. F 1950
- BILLINGS, Edward G. (M.D.)**
1820 High Street
Denver 6, Colo. F 1940
- BISSIRI, Gerald R. Mr.**
4410 West Avenue 42
Los Angeles 65, Cal. A 1959
- BLAU, Theodore H. (Ph.D.)**
213 E. Davis Blvd.
Tampa 6, Fla. A 1955
F 1956
- BLESSING, Harold D. (Ph.D.)**
29 Grindon Dr.
Wilmington 8, Del. A 1953
- BLUM, Gerald S. (Ph.D.)**
2641 Geddes Ave.
Ann Arbor, Mich. F 1961
- BLUMSTEIN, Mrs. Molly G.**
5219 Wayne Avenue
Philadelphia 44, Pa. A 1948
- BOGARDUS, Helen (Ed.D.)**
1212 Fifth Ave. No. #301
Seattle 9, Washington A 1949
- BOLGAR, Hedda (Ph.D.)**
Psychiatric and
Psychosomatic
Research Institute,
Mt. Sinai Hospital
8720 Beverly Blvd.
Los Angeles 48, Calif. F 1956
- BONDEL, Gertrude (Ph.D.)**
2049 McGraw Ave. St. Aff. 1953
New York 62, N.Y. A 1954
- BORTREE, David**
P.O. Box 871
Kansas City, Mo. A 1961
- BOSNER, Mrs. Jane Potter**
58 Ridge Road
Rumson, N.J. St. Aff. 1957
- BOURKE, William T. (Ph.D.)**
135 East 73rd St.
New York 21, N.Y. A 1950
- BOWDLEAR, Charles M. (Ph.D.)**
Psychology Services
VA Hospital
Sepulveda, Calif. A 1957
- BOWERS, Scott T. (Ed.Dr.)**
870 Reibold Building
Dayton 2, Ohio A 1956
- BRADWAY, Katherine P. (Ph.D.)**
41 Monte Mar Drive
Sausalito, Calif. A 1940
F 1949
- BRAMWELL, Paul F.**
4638 Locust Lane
Salt Lake City, Utah A 1957
- BRANDT, Rudolph J. (Ph.D.)**
435 N. Roxbury Drive
Beverly Hills, Calif. A 1950
- BRAUN, Mrs. Roslyn R.**
170-15 Highland Avenue
Jamaica Estates 32, N.Y. A 1949
- BREWER, Mrs. Hermione**
359 E. Baltimore
Greencastle, Pa. A 1951
- BREWER, Paul W. (Ph.D.)**
1301 E. Palo Verde Dr.
Phoenix 14, Ariz. A 1961
- BRICKLIN, Mrs. Patricia McIntosh**
78 Hartwood Road
Levittown, Pa. A 1958
- BRISKIN, Gerald J. (Ph.D.)**
1594 Villa Road
Birmingham, Mich. A 1961
- BRODIE, Mrs. Dorothy B.**
4833 Tenth Ave. N.
St. Petersburg, Fla. A 1943
F 1950
- BRODY, B. Abraham (Ph.D.)**
29 Heatherbloom Road
White Plains, N.Y. A 1952
- BRODY, Claire M. (Ph.D.)**
73 Ivy Lane
Englewood, N.J. A 1955
- BRODY, Gertrude Gillenson**
29 Heatherbloom Road
White Plains, N.Y. A 1948
- BROOMHEAD, Elizabeth**
5350 42nd Place, N.W.
Washington 15, D.C. A 1944
- BROSIN, Henry W. (M.D.)**
University of Pittsburgh
Western Psychiatric Institute
3811 O'Hara Street
Pittsburgh 13, Pa. F 1940
- BROWER, Daniel (Ph.D.)**
300 N. Mountain Ave.
Upper Montclair, N.J. A 1943
F 1954
- BROWER, Mrs. Judith F.**
300 N. Mountain Avenue
Upper Montclair, N.J. A 1948
- BROWN, Fred (Ph.D.)**
Mt. Sinai Hospital
Fifth Ave. and 100th St. A 1948
New York, N.Y. F 1950
- BROWNELL, Rosa Parsons**
2725 Barnson Pl.
San Diego 3, Calif. A 1947
- BROWNFAIN, John J. (Ph.D.)**
21174 Greenview Ave.
Southfield, Mich. A 1954
- BROZOVICH, Stanley M.**
449 E. Pine
Altadena, Calif. A 1953
- BRUCE, Martin M. (Ph.D.)**
340 Oxford Rd.
New Rochelle, N.Y. A 1952
- BRUNDSCHWIG, Lily (Ph.D.)**
3287 Washington Blvd.
Cleveland Hts. 18, Ohio A 1953
- BRY, Mrs. Mae G.**
59 W. 12th Street
New York 11, N.Y. A 1954
- BUEGEL, Hermann F.**
Dept. of Psychology
University of No. Dakota
Grand Forks, No. Dakota A 1961
- BUHLER, Charlotte (Ph.D.)**
1127 N. Sweetzer
Los Angeles 46, Calif. A 1943
F 1951
- BURCHARD, Edward M. L. (Ph.D.)**
1230 Park Ave.
New York 28, N.Y. F 1940
- BURDICK, Mrs. Jeanne H.**
2236 Edgewater Ter.
Topeka, Kans. Aff. 1957
- BURGEMESTER, Bessie B. (Ph.D.)**
1175 York Avenue
New York 21, N.Y. A 1942
F 1947
- CALABRESI, Renata A. (Ph.D.)**
360 Central Park West
New York 25, N.Y. F 1950
- CALIGOR, Leopold (Ph.D.)**
175 Riverside Drive
New York 24, N.Y. A 1952
- CALVERT, Margaret**
73510 Mountain St., Apt. 1
Montreal, Quebec, Can. A 1952
- CANTER, Aaron H. (Ph.D.)**
2939 N. 47th Street
Phoenix 18, Ariz. A 1949
F 1952
- CAPALDI, Mrs. Betty F.**
86 Westview Rd., Apt. 5
Wilmington 2, Del. A 1959
- CAPELL, Martin D. (Ph.D.)**
18258 Fremont St.
Livonia, Michigan A 1958
- CARO, Mrs. Elizabeth R.**
2014 Grove Avenue
Richmond, Va. A 1940
- CARPENTER, Kenneth E. (Ph.D.)**
1164 Kingston Road,
Kingston, R. I. A 1953
- CARR, Arthur C. (Ph.D.)**
Psychiatric Institute
722 West 168th Street
New York 32, N.Y. A 1953
- CARROLL, Clara**
Bureau of Child Guidance
80 Lafayette, 11th Floor
New York 13, N.Y. F 1940
- CARSON, Marjorie**
Children's Aid Soc. of
Metropolitan Toronto
33 Charles Street East
Toronto 5, Ont., Can. A 1952
- CARTER, Linda Louise**
1115 Wertland Street
Charlottesville, Va. A 1949
- CARTWRIGHT, Robert W.**
22364 MacFarlane Dr.
Woodland Hills, (Ph.D.)
Calif. A 1952
- CASSEL, Russell N. (Ed.D.)**
501 No. Poppy St.
Lompoc, Calif. A 1954
F 1955
- CEASE, Eugene**
Box 234
Warren State Hospital
Warren, Pa. A 1951
- CENTERS, Louise Van Core (Ph.D.)**
3676 Mandeville Canyon Rd.
Los Angeles 49, Calif. A 1958
- CHAMBERLAIN, Allan B.**
21 Academy St.
Cambridge, N.Y. A 1961
- CHARNY, Israel W. (Ph.D.)**
Oakbourne Hospital
1030 S. Concord Rd.
West Chester, Pa. A 1957
- CHAYKIN, Albert**
Guidance Center
University of Miami
Coral Gables, Fla. A 1956

- CHU, Thomas W.
770 West End Avenue
New York 25, N. Y. A 1955
- CICCARELLO, Miss Jennie
Psychology Dept.
Univ. of Portland
St. Aff. 1956
Portland 3, Oregon A 1958
- CLARK, W. Donald
8 Rosalind Rd.
Frenton 8, N.J.
St. Aff. 1960
- CLAUSS, Helen O.
23 Glen Street
Chambersburg, Pa. A 1951
- CLERK, Mrs. Gabrielle Brunet
249 Lockhart
Town of Mt. Royal
Quebec, Can. A 1949
- COHEN, Mrs. Mathilde Weill
46 E. 91st Street
New York 28, N. Y. A 1942
- COHN, Mrs. Ruth C.
159 Liberty Road
Englewood, N. J. A 1946
- COLE, Joseph Carl (Ph.D.)
14034 S. Pioneer Blvd.
Norwalk, Calif. A 1949
- COLM, Hanna (Ph.D.)
3 Overhill Road
Falls Church, Va. A 1944
- COLVIN, Ralph (Ph.D.)
Astor Home for Children
Rhinebeck, N.Y. A 1956
- CONANT, James C. (Ph.D.)
2509 Road 64
Pasco, Wash. A 1961
- COOK, Philip H. (Ph.D.)
Department of Labour and
National Service
Swanston Street
Melbourne c.1
Victoria, Australia A 1941
F 1949
- COOPER, Miss Gertrude A 1961
3004 No. Stuart St.
Arlington 7, Va. A 1961
- COPEL, Sidney L.
5606 Wyndale Avenue
Philadelphia 31, Pa. A 1956
- COTTES, Mlle. S.
213 Rue d'Endoume
Marseille, France F 1961
- COUGHLIN, Mrs. Anne Malone
Dept. of Psychiatry
Montreal General Hospital
Montreal, P. Q. St. Aff. 1954
- *COWIN, Marion
433 West 21st Street
New York 11, N. Y. F 1940
- COX, Rachel Dunaway (Ph.D.)
503 Walnut Lane
Swarthmore, Pa. A 1950
F 1952
- CRADDICK, Ray (Ph.D.)
Box 915
University Park, N. Mex.
A 1960
- CRASILNECK, Harold B.
(Ph.D.)
Dallas Neurological Clinic
Medical Tower
712 No. Washington
Dallas 10, Texas A 1956
- CRILE, Mrs. Mary
Big Sur, Calif. A 1945
- CROVETTO, Lorraine
703 Carondelet Street
New Orleans 12, La. A 1953
- CRUMPTON, Evelyn (Ph.D.)
1451 1/2 Barry Avenue
Los Angeles 25, Calif. A 1955
- CRYNIS, Gerd M. (Ph.D.)
C. G. Jung Inst.
Gemeindestrasse 27
Zurich 32, Switzerland A 1960
- CUMMINGS, C. Peter (Ph.D.)
555 Weadley Road
Strafford, Pa. A 1954
- CUNNINGHAM, Mrs. Cornelia
84 E. Moreland Ave.
Philadelphia 18, Pa. A 1950
- DaCUNHA, Dr. M. C.
Opposite the Dargah
Cadell Road
Mahem, Bombay 16, India
A 1961
- DANA, Richard H. (Ph.D.)
Dept. of Philosophy &
Psychology
West Virginia Univ.
Morgantown, W. Va. A 1956
- D'ANGELO, Rita Y.
2604 University Avenue
Bronx 68, N. Y. A 1955
- DAVENPORT, Beverly (Ph.D.)
7701 Macaw Lane
San Diego 11, Calif. A 1949
- DAVID, Dr. Charlotte
1630 N.E. 50th Ave.
Portland 13, Oregon A 1959
- DAVID, Henry P. (Ph.D.)
New Jersey State
Dept. of Inst. & Agencies
State Office Building
135 W. Hanover St.
Trenton 25, N.J. A 1953
- DAVIDS, Anthony (Ph.D.)
Dept. of Psychology
Brown University
Providence, R.I. F 1961
- DAVIDSON, Alene (Ph.D.)
7658 175th St.
Flushing 66, N.Y. A 1953
- *DAVIDSON, Helen H. (Ph.D.)
90 La Salle St.
New York, N.Y. F 1940
- DAVIS, John A. (Ph.D.)
2444 Archwood Drive
Dayton 6, Ohio A 1955
- DAVISON, Arthur H. (Ph.D.)
Psychology Dept.
Montana State Hosp.
Warm Springs, Mont. F 1953
- DEAN, Sidney I. (Ph.D.)
1630 Liholiho St.
Honolulu 14, Hawaii A 1958
- DECKER, Robert J.
108 Hampden Ave.
Narberth, Pa. A 1961
- DeDIAZ, Mrs. Nelly
Juventino Rosas
#70 Col Guadalupe Inn
Mexico, D.F., 20 St. Aff. 1961
- DeGOVIA, Guillermo
Box 25333
University of Mexico
Mexico, D.F. St. Aff. 1961
- DeLENERO, Mrs. Estela
Avenue Cuauhtemoc
#719-203
Mexico, D.F. St. Aff. 1961
- DE MARTINO, Hugo A.
2823 Walker Drive, RFD 2A
Peekskill, N.Y. A 1958
- DERI, Mrs. Susan K.
235 W. 76th Street
New York 23, N. Y. A 1948
F 1950
- DERNER, Gordon F. (Ph.D.)
Department of Psychology
Adelphi College
Garden City, N. Y. A 1949
F 1951
- DE VAULT, Mrs. Barbara Allen
Division of Legal Medicine
33 Broad St.
Boston, Mass. A 1953
- DE VAULT, Helen C.
2312 Via Pinale
Palos Verdes Estates, Calif.
A 1950
- DE VOS, George (Ph.D.)
Lecturer in Social Welfare
Univ. of California
Berkeley 4, Calif. F 1957
- DIAMOND, Mrs. Florence
135 Sierra View Road
Pasadena 2, Calif. A 1950
- DIANA, Pearl Butler (Ph.D.)
4111 Hillcrest Rd.
Richmond 23, Va. A 1949
F 1951
- DICK, Miss Frances
209 Lincoln Place
Brooklyn 17, N.Y. A 1958
- DIEPPA, Jorge J. (Ph.D.)
P.O. Box 1065
San Juan, P.R. A 1958
- DINGMAN, Paul R. (Ph.D.)
1206 Pleasant
Des Moines 14, Iowa A 1950
- DOAK, Mrs. James
High St.
Rockport, Maine A 1953
- DOMINGUEZ, Kathryn (Ph.D.)
126 Tyson Road
Newtown Square, Pa. A 1943
- DONOGHUE, John R.
Psychology Dept.
Lincoln St. Hospital
Lincoln 10, Nebr. St. Aff. 1959
- DORKEN, Herbert, jr. (Ph.D.)
Community Mental Health
Services
Dept. of Public Welfare
658 Cedar Ave.
St. Paul 1, Minn. A 1949
F 1951
- DOUGHERTY, Mrs. Margaret R.
1804 Roselynn Avenue
Scranton 10, Pa. A 1944
- DOWLEN, Caroline (Ph.D.)
Box 252
Elko, Nev. A 1956
- DRYSELIOUS, Harold
601 S. Gramercy Place
Los Angeles, Calif. A 1950
- DUDEK, Mrs. Stephanie Z.
(Ph.D.)
781 Beatty Ave.
Montreal 19, Que., Can. A 1949
F 1961
- DUE, Floyd O. (M.D.)
370 29th Street
Oakland 9, Calif. A 1943
- DUNLAP, Dorothy
Agnew State Hospital
San Jose 14, Calif. A 1954
- DUNN, Michael B. (Ph.D.)
405 No. Bedford Drive
Beverly Hills, Calif. A 1941
F 1943
- EBER, Milton (Ph.D.)
Inst. Jackson
Memorial Hospital
Miami, Fla. A 1961
- EGLASH, Mrs. Evelyn
Detroit Inst. of Technology
131 East Adams
Detroit 26, Mich. A 1953
- EIDUSON, Dr. Bernice T.
(Ph.D.)
9760 W. Pico Blvd.
Los Angeles 35, Calif. A 1949
- EISNER, Betty Grover (Ph.D.)
1334 Westwood Blvd.
St. Aff. 1955
Los Angeles 24, Calif. A 1961
- ELDRD, Donald M.
Psychology Department
Vermont State Hospital
Waterbury, Vt. A 1948

- ELIZUR, Abraham (Ph.D.)
6 Tel Hai Street
Tel Aviv, Israel A 1949
- ELLIS, Albert (Ph.D.)
Pare Vendome
353 W. 56th Street
New York 19, N.Y. A 1950
- ENOCHS, Neil
366 Marie Avenue S.E. A.E. 1954
Los Angeles 42, Calif.
- EPSTEIN, Hans L. (Ph.D.)
222 W. 176th Street
New York 33, N.Y. A 1944
- ERICSON, Mrs. Helen
11844 E. Deana Street
El Monte, Calif. A.E. 1954
- ERON, Leonard D. (Ph.D.)
Rip Van Winkle Foundation
4-4 Warren Street
Hudson, N.Y. F 1953
- ESTRADA, Mrs. Carol Griffin
4256 Maryland Ave.
St. Louis 8, Mo. A 1951
- EVANS, John T. (Ph.D.)
85 Otis Street
Newtonville 60, Mass. F 1957
- EVANS, Ray B.
1575 View Site Drive
Los Angeles 46, Calif. A 1954
- EVERETT, Evelyn G. (Ph.D.)
Box 51
Imola, Calif. A 1955
- EKKER, John E. Jr.
Bur. of Testing & Research
De Pauw University A 1957
Greencastle, Indiana F 1961
- FABRIKANT, Benjamin (Ph.D.)
248 Devon Road A 1959
Westwood, N.J. F 1961
- FARBROW, Norman L. (Ph.D.)
4211 Holly Knoll Drive
Los Angeles 27, Calif. A 1949
- FARLEY, Jane (Ph.D.)
Palmb Hearing Inst.
Loyola University
6425 N. Sheridan
Chicago 26, Ill. A 1960
- FATFERN, Hanna F. (Ph.D.)
1 Sheridan Square A 1943
New York 16, N.Y. F 1946
- FEHRENBACH, Dr. Alice
222 S. Josephine St.
Denver, Colo. A 1951
- FEIFFEL, Herman (Ph.D.)
VA Outpatient Clinic
1031 S. Broadway A 1943
Los Angeles 15, Calif. F 1956
- FEIN, Leah Gold (Ph.D.)
32 Urban St.
Stanford, Conn. H.F. 1961
- FERN, Leah Gold F 1961
32 Urban St.
Stanford, Conn.
- FEINBERG, Henry
15886 La Salle A 1949
Detroit, Mich.
- FELDBERG, Theodore M. (M.D.)
11 E. Chase Street
Baltimore 2, Md. A 1944
- FELDMAN, Dorothy A. (Ph.D.)
Medical Arts Bldg.
Pittsburgh 13, Pa. A 1952
- FELDMAN, Irving (Ph.D.)
80 Poplar Ave.
W. Long Branch, N.J. A 1953
- FELZER, Stanton B., (Ph.D.)
340 Ainslie Rd. A 1954
Huntington Valley, Pa. F 1961
- FERGUSON, Kingsley G. (Ph.D.)
Psychology Department
Westminster Hospital
London, Ontario, Can A 1954
- FERNANDEZ, Miss Rosa
Ortega 232 Coysacon
21 Mexico, D.F. St. Aff. 1961
- FERRACUTI, Franco (Prof.)
Via Ugo Balzani 57
Roma, Italy A 1954
- FEUERBURGH, Joseph (Ph.D.)
15 Stuyvesant Oval
New York 9, N.Y. A 1957
- FICHMAN, Lionel L. (Ph.D.)
241 Church Lane
St. Aff. 1954
Los Angeles 49, Calif. A 1957
- FILMFR-BENNETT, Gordon
Winnebago State Hospital
Winnebago, Wis. A 1954
F 1956
- FINE, Reuben (Ph.D.)
225 W. 86th Street
New York 24, N.Y. A 1949
F 1954
- FINN, Michael H. P. (Ph.D.)
7820 Ellenham Ave. A 1954
Ruxton, Maryland F 1958
- FISCHER, Liselotte K. (Ph.D.)
219 Bryant Street
Buffalo 22, N.Y. A 1949
- FISHER, Emanuel (Ph.D.)
274 First Ave.
New York 9, N.Y. A 1960
- FITE, June Harris (Ph.D.)
Hunter College
Educational Clinic
695 Park Ave.
New York 21, N.Y. A 1941
- FONT, Marion McKenzie
627 S. Carrollton Ave. A 1942
New Orleans 13, La. F 1947
- FORER, Bertram R. (Ph.D.)
2170 Live Oak Drive E. A 1949
Los Angeles 28, Calif. F 1951
- FORER, Lucille K. (Ph.D.)
2170 Live Oak Drive E.
Los Angeles 28, Calif. A 1953
- FORREST, Dr. Carol W.
Opfergasse, 18
Vienna 4, Austria A 1951
- FORTIER, Robert H. (Ph.D.)
Chief Psychologist
VA Mental Health Clinic
Indianapolis, Ind. A 1956
- *FOSBERG, Irving A. (Ph.D.)
7831 Plum St. A 1940
New Orleans 10, La. F 1949
- FOSTER, Austin (Ph.D.)
Suite 209, Doctor's Bldg.
800 Fifth Ave. A 1950
Fort Worth, Texas F 1955
- FRAMO, James L., Jr. (Ph.D.)
342 Merion Rd. A 1955
Merion, Pa. F 1957
- FRANCOEUR, Thomas A.
1070 Crevier Avenue (Ph.D.)
Ville St. Laurent
Prov. Que., Canada Aff. 1954
- FRANK, Lawrence K.
25 Clark Street
Belmont, Mass. H.M. 1954
- FRANKEL, Esther B. (Ph.D.)
Child Guidance Clinic
Children's Hospital
San Francisco 18, Cal. A 1953
- FREAR, Edgar
Montrose, Pa. A 1950
- FRESCHI, Vincent J. (Ph.D.)
14 Penn Lane
West Chester, Pa. A 1960
- FREY, Mrs. Harriet K.
59 Francisco Avenue
West Caldwell, N.J. A 1953
- FRIEDMAN, Mrs. Gladys Miller
29575 So. Woodland Blvd.
Pepper Pike 24, Ohio A 1949
- FRIEDMAN, Howard (Ph.D.)
316 Southfield Drive
Havetteville, N.Y. A 1951
- FRIEDMAN, Ira (Ph.D.)
29575 So. Woodland Blvd.
Pepper Pike 24, Ohio A 1954
F 1958
- FRIEND, Mrs. Jeannette G.
16 Greenough Circle
Brookline 46, Mass. A 1949
- FRISCH, Paul Z. (Ph.D.)
Psychology Department
Adelphi College
Garden City, N.Y. F 1956
- *FROMM, Erika O. (Ph.D.)
5717 S. Kenwood
Chicago 37, Ill. A 1940
- FROSTIG, Marianne (Ph.D.)
7257 Melrose Ave.
Los Angeles 46, Calif. A 1956
- FRY, Franklin D.
1724 Wyoming Avenue
Forty Fort
Wilkes-Barre, Pa. A 1952
- FRY, Mrs. Martha O.
1724 Wyoming Avenue
Forty Fort
Wilkes-Barre, Pa. A 1952
- FUCHS, Arnold
86-06 35th Ave.
Jackson Heights 72, N.Y.
St. Aff. 1960
- FUCHSMAN, Seymour H.
140 E. 40th St.
New York 16, N.Y. A 1944
- FULLER, Gerald B. (Ph.D.)
Director of Research
Willmar State Hospital
Willmar, Minn. A 1961
- FURCHNER, Robert (Ph.D.)
5107 N. Willamette Blvd.
Portland, Ore. A 1961
- GALLICO, Mrs. Margaret Wilson
Univ. of Dayton
Dayton 9, Ohio A 1957
- GARFIELD, S. L. (Ph.D.)
Nebraska Psychiatric Inst.
602 South 44th Ave.
Omaha, Neb. F 1959
- GASOREK, Miss Kathryn
29 West Henry St.
Linden, New Jersey A 1949
- GASTON, Charles O.
University of Texas
Medical Branch St. Aff. 1955
Galveston, Texas A 1957
- *GAUDET, E. Louise (Ph.D.)
211 East 53rd St.
New York 22, N.Y. F 1940
- GAUDET, Frederick J. (Ph.D.)
210 W. 70th St. A 1949
New York 23, N.Y. F 1958
- GAVIDIA, Alvaro Villar, M.D.
Avenida 42 #15-71
Bogota, Columbia A 1957
- GEIL, George A.
219 E. Seminole
Springfield, Mo. A 1943
- *GERING, Mrs. Evelyn E.
18100 Karen Drive
Tarzana, Calif. A 1940
- GERSTEN, Rev. Charles
Via Coeli Monastery (Ph.D.)
Jemez Springs, N.M. A 1949
- GESSNER, Alan (Ph.D.)
1806 New Jersey Road
Lakeland, Fla. A 1961
- GETOFF, Louis
317 West 90th St.
New York 24, N.Y. A 1956

- GIBSON, Robert I. (Ph.D.)
R.D. 5,
Norwich, Conn. A 1960
- GILBERT, Raymond R.
52 Halifax Street
Boston 30, Mass. A 1951
- GILLMAN, Mrs. Etta C.
16 Stevenson Avenue
Hartsdale, N. Y. A 1944
- GINANDES, Mrs. Janet
1150 Fifth Ave.
New York, N.Y. A 1957
- GLADFELTER, John (Ph.D.)
Dept. of Psychiatry
Southwestern Medical School
Dallas, Texas A 1958
- GLASS, Blanche (Ph.D.)
35 E. 85th St.
New York 28, N.Y. A 1955
- GOLDBERGER, Leo
Research Center for
Mental Health
New York University
21 Washington Place
New York 3, N.Y. A 1956
- GOLDBLOOM, Betty M. (Ph.D.)
5642 Darlington Rd.
Pittsburgh 17, Pa. A 1952
- GOLDEN, Doris Schulman
10 Downing Street
New York 14, N. Y. A 1948
- GOLDFARB, William (M.D.)
530 West End Avenue
New York 31, N. Y. F 1944
- GOLDSTEIN, Dr. Fred J.
353 S. Maple Drive
Beverly Hills, Calif. A 1956
- GONDOR, Mrs. Lily H.
320 East 57th Street
New York 22, N. Y. F 1952
- GOODMAN, Mrs. Beverly
Schillinger
143-50 Hoover Ave.
Jamaica 35, N.Y. A 1956
- GOODMAN, Morris (Ph.D.)
2130 Millburn Ave.
Maplewood, N. J. A 1953
- GOODMAN, Mrs. Paya
165 E. 179th St.
Bronx, N.Y. A 1961
- GOODNICK, Benjamin (Ph.D.)
Administration Building
Parkway and 21st
Philadelphia 5, Pa. A 1956
- COOLISHIAN, Harold A. (Ph.D.)
1008 Camp Circle West A 1952
La Marque, Texas F 1957
- GORDON, Thelma
307 W. 11th Street
New York 14, N. Y. A 1951
- GOTTLIEB, Mrs. Sophie B.
225 W. 86th Street
New York 24, N. Y. A 1943
- GRAHAM, Virginia T. (Ph.D.)
General Hospital
(Central Clinic)
Cincinnati 29, Ohio A 1953
- GRASSI, Joseph R.
Grayslyn Hospital
Winston-Salem, N. C. A 1942
- GRAVES, Winifred S. (Ph.D.)
4242 Cornelius Ave.
Indianapolis 8, Ind. F 1951
- GRAVITZ, Melvin A. (Ph.D.)
2025 Eye St., N.W.
Washington, D.C. F 1961
- GRAYSON, Harry M.
403 S. Bundy Dr.
Los Angeles, Calif. A 1951
- GREENBERG, Nathan
2820 Darlington Pl., Apt. 30
Montreal, Que., Can. A 1959
- GREENE, Janet S. (Ph.D.)
65 E. 76th Street
New York 21, N. Y. A 1953
- GREENSTADT, William M.
35 E. 30th Street St. Aff. 1954
New York, N. Y. A 1953
- GREENSTEIN, Mr. Jules M.
20 Rector
Metuchen, N.J. A 1960
- GRIER, Mary E. (Ph.D.)
2243 Harcourt Drive
Cleveland 6, Ohio A 1956
- GROFF, Marie L. (Ph.D.)
Holly Knoll, Arbutus Drive
Rte. 3, Annapolis, Md. A 1952
- GROSSMAN, Searles A. (Ph.D.)
4004 Coleridge Road A 1951
Wilmington 2, Del. F 1954
- GUERTIN, Wilson H. (Ph.D.)
College of Nursing
Univ. of Florida A 1950
Gainesville, Fla. F 1953
- GUIORA, A. Zeev (Ph.D.)
4 Pines Street
Rehovoth, Israel A 1957
- GUNDLACH, Ralph (Ph.D.)
162 East 80th Street
New York 21, N. Y. A 1951
- GURVICH, Mrs. Bernice M.
251 Willis Avenue
Hawthorne, N. Y. A 1950
- GURVITZ, Milton S. (Ph.D.)
54 Gateway Dr.
Great Neck, N.Y. F 1951
- GUY, William
2940 Ellsworth Rd.
Ypsilanti, Mich. A 1953
- GUZE, Mrs. Vivian S.
299 Clinton St.
Newark, N.J. A 1960
- HABER, Wm. B. (Dr.)
275 Central Park West
New York 24, N. Y. A 1953
- *HALLOW, William C. (Ph.D.)
422 Chestnut St.
Lebanon, Pa. A 1940
- *HALLOWELL, A. Irving (Ph.D.)
Box 14, Bennett Hall
Univ. of Pennsylvania A 1940
Philadelphia 4, Pa. F 1944
- HALLOWELL, Dorothy K.
3318 Midvale Avenue (Ph.D.)
Philadelphia 29, Pa. A 1947
- HALPERIN, Sidney L. (Ph.D.)
1710 Makiki St.
Apt. 906
Honolulu 14, Hawaii A 1949
- HALPERN, Esther (Ph.D.)
1400 Pine Ave. W.
Apt. 803
Montreal 25, Que., Can. St. Aff. 1954
- HALPERN, Florence (Ph.D.)
University Hospital
303 E. 20th St.
New York City, N.Y. F 1959
- HAMILTON, Sidney F.
Box 5314 NT Station
Denton, Texas A 1961
- HAMMER, Emanuel F. (Ph.D.)
685 West End Avenue A 1953
New York 25, N.Y. F 1959
- HAND, Dr. Mary Ella
326 Fifth St.
Ann Arbor, Mich. A 1948
- HANDEL, Gerald
Social Research, Inc.
145 East Ohio Street
Chicago 11, Ill. A 1954
- HANSEN, Irvin (Ph.D.)
360 No. Bedford Drive
Beverly Hills, Calif. A 1960
- HARMES, John M.
240 Smith Road
Manchester, N.H. A 1957
- HARRIS, Albert J. (Ph.D.)
Educational Clinic
Queens College
Flushing, N. Y. A 1951
- HARRIS, Robert A. (Ph.D.)
Austin Riggs Center Inc.
Stockbridge, Mass. A 1954
- HARRIS, Robert E. (Ph.D.)
The Langley Porter Clinic
University of California
Medical Center
San Francisco 22, Calif. A 1948
- HARRIS, William W.
2300 Grand Concourse
Bronx 58, N.Y. A 1949
- *HARROWER, Molly R. (Ph.D.)
55 E. 86th Street
New York 28, N. Y. F 1940
- HAWORTH, Mary R. (Ph.D.)
Nebraska Psychiatric Inst.
602 S. Fourth Ave.
Omaha 5, Neb. A 1959
- HAYS, Miss Betty
305 S. Harvard Blvd
Los Angeles 5, Calif. A 1949
- HEATH, Douglas (Ph.D.)
48 Wentworth Lane
Rosemont, Pa. A 1956
- HEBERT, Bernard
4368 Girard
Montreal 28, Canada A 1955
- HEISLER, Verda (Ph.D.)
2306 - 8th Avenue
San Diego, Calif. A 1951
- HELLFRSBERG, Elisabeth F.
641 Whitney Avenue (Ph.D.)
New Haven, Conn. A 1919
- HEMMENDINGER, Larry
58 Judd St.
Fairfield, Conn. (Ph.D.)
A 1950
- *HENRY, William E. (Ph.D.)
5835 Kimbark Avenue A 1948
Chicago, Ill. F 1956
- HERMAN, Jack L.
2546 Horace Court
Bellmore, L.I., N.Y. A 1961
- HERNESS, Mrs. Christina
Amherst H. Wilder Child
Guidance Clinic
670 Marshall Avenue
St. Paul 4, Minn. A 1952
- *HERTZ, Marguerite R. (Ph.D.)
2835 Drummond Road
Shaker Heights, Ohio F 1940
- *HERTZMAN, Max (Ph.D.)
Dept. of Psychology
College of the
City of New York A 1940
140th and Convent Ave.
New York, New York F 1946
- HIGGINSON, Gordon K. (Ph.D.)
6040 N. Montana A 1954
Portland, Ore. F 1959
- *HILDEN, Arnold H. (Ph.D.)
628 Clark Avenue A 1940
Webster Groves 19, Mo. F 1943
- HILKEVITCH, Rhea R. (Ph.D.)
1550 Oak Avenue
Evanston, Ill. A 1954
- HILLSON, Joseph (Ph.D.)
Norfolk State Hospital
Norfolk, Neb. A 1956
- HILTMANN, Hildegard, Dr.
Lehrstuhl für
Angewandte Psychologie
an der Universität
Freiburg im Breisgau
Peterhof, Peterstr. 1
Germany F 1957

- HIMELSTEIN, Philip. (Ph.D.)**
Dept. of Psych.
Univ. of Arkansas
Fayetteville, Ark. A 1956
- HINDS, Edith A. (Ph.D.)**
1239 Lincoln Place
Brooklyn, New York A 1960
- *HIRNING, L. C. (M.D.)**
R F.D. No. 1, Box 180
Katonah, N.Y. F 1940
- HIRSCH, Mrs. Janet F.**
67 49 C 192nd Street
Fresh Meadows 65, N.Y. A 1948
- HOCH, Erasmus L. (Ph.D.)**
952 Anneton Road
Bethesda 14, Md. A 1954
- HOLANCHOKE, George**
Box 132
Claymont, Dela. A 1967
- HOLMES, Frances B. (Ph.D.)**
Whetstone Rd. R.D. 2
Hartwinton, Conn. A 1950
- HOLODNAK, Helen Barbara**
31-38-38th Street
Astoria 3, L.I., N.Y. A 1949
- HOLT, James M. (Ph.D.)**
5554 Littlebow Road
Palos Verdes Estates, Calif. A 1956
- HOLTZMAN, Wayne (Ph.D.)**
Hogg Foundation
University of Texas
Austin, Texas F 1959
- HOLZBERG, Jules D. (Ph.D.)**
Box 351
Middletown, Conn. F 1954
- HOOKER, Mrs. Evelyn (Ph.D.)**
400 S. Saltair Ave.
Los Angeles 49, Calif. F 1958
- HORICK, Reuben S. (Ph.D.)**
4004 N. Stuart Street
Arlington 7, Va. F 1961
- HORN, Daniel (Ph.D.)**
521 West 57th St.
New York 19, N.Y. F 1959
- HOSHINO, Akira**
Department of Psychology
International Christian Univ.
Mitaka Tokyo, Japan A 1951
- HOWARD, J. W. (Ph.D.)**
Route 2
Rigaud, P.Q. Canada A 1954
- HOWARD, J. Stephen**
1132 So. Keniston Ave.
Aff. 1954
Los Angeles 19, Calif. A 1958
- HOWLAND, Allan O.**
2409 1/2 Panama St.
Philadelphia 3, Pa. A 1951
- HUBER, Jack T. (Ph.D.)**
310 East 55th St.
New York 22, N.Y. F 1958
- HUGHES, Robert M. (Ph.D.)**
Suite 101
849 Peachtree St., N.E. A 1944
Atlanta 8, Ga. F 1954
- HUTT, Max L. (Ph.D.)**
2114 Vinewood Blvd.
Ann Arbor, Mich. A 1947
F 1952
- IMRE, Paul**
2111 Drummond Rd.
Catonville 28, Md. A 1954
- INMAN, John M.**
160 Tamalpais Road
Berkeley 8, Calif. A 1945
- INNES-SMITH, Dr. J.**
195 Laurier Ave., East
Ottawa 2, Ont., Can. A 1959
- ISAACS, Mark, Dr.**
Chief of Psych. Services
Rosewood State Training Sch.
Owning Mills, Md. A 1959
- IVERSON, Norman E. (Ph.D.)**
North Central Medical Bldg.
2021 North Central Ave.,
Phoenix, Ariz. A 1956
- IVES, Margaret (Ph.D.)**
5th Elizabeth Hospital
Washington 20, D.C. A 1955
F 1955
- JACOBS, Martin E. (Ph.D.)**
999 Central Ave.
Wooner, L.I., N.Y. A 1955
- JEFFRIES, Mrs. Helen**
14 East Sixth Street
Media, Pa. A 1956
- JOEL, Walther (Ph.D.)**
9629 Brighton Way
Beverly Hills, Calif. A 1946
F 1950
- JOHNSON, Elizabeth Z.**
(Ed.D.)
366 Garden Rd.
Lexington, Ky. F 1956
- JOHNSON, Richard B.**
110 Waverly Place
New York 11, N.Y. A 1953
- JOHNSON, Theresa**
229 S. Maple Drive
Beverly Hills, Calif. A 1949
- JOHNSON, Thomas F. (Dr.)**
3 Richie Lane
Yardley, Pa. A 1957
- JONES, Marshall R. (Ph.D.)**
Univ. of Nebraska, Psychology
Lincoln, Neb. F 1961
- JORTNER, Sidney**
1425 Brooklyn Avenue
Brooklyn 10, N.Y. A 1959
- JOSEPH, Alice (M.D.)**
114 E. 71st St.
New York 21, N.Y. A 1914
- *JUNKEN, Elizabeth M. (Ph.D.)**
468 Lydecker Street
Englewood, N. J. A 1940
- KADINSKY, D.**
8 P. Smolenski Street
Tel Aviv, Israel A 1946
- KADIS, Mrs. Asya L.**
1060 Park Avenue
New York 28, N.Y. A 1944
- KAHN, David F. (Ph.D.)**
Lexington School for Deaf
904 Lexington Avenue
New York 21, N.Y. A 1953
- KAHN, Marvin W. (Ph.D.)**
University of Colorado
Medical School
4200 E. 9th Avenue
Denver, Colo. A 1956
F 1959
- KALINKOWITZ, Bernard N.**
Graduate School of
Arts and Science
New York University
Washington Square
New York, N.Y. A 1954
- KALIS, Betty Lee (Ph.D.)**
98 Levant St.
San Francisco 14, Calif. A 1956
- KAPIT, Milton E.**
1 W. 85th Street
New York 21, N.Y. A 1950
- KAPLAN, Donald M. (Ph.D.)**
41 Fifth Ave.
New York 3, N.Y. A 1961
- KAPLAN, Herbert**
Patricia Avenue
Fishkill, N.Y. A 1949
- KAPLAN, Norman (Ph.D.)**
302 W. 35th St.
Savannah, Ga. A 1949
- KAPLIN, Bert (Ph.D.)**
Dept. of Psychology
Univ. of Kansas
Lawrence, Kansas A 1958
- KARSON, Samuel (Ph.D.)**
Dade County Child Guidance
Clinic
1350 N.W. 14th St.
Miami, Florida F 1957
- KASS, Walter (Ph.D.)**
4 Farley Road
Scarsdale, New York F 1955
- KATAGUCHI, Yasufumi**
Natl. Inst. of Mental Health
Kohadai Ichikawa
Chiba, Japan A 1958
F 1961
- KATES, Solis L. (Ph.D.)**
University of Massachusetts
Amherst, Mass. A 1949
- KATZ, Mrs. Florine**
67 East 82nd Street
New York 28, N.Y. A 1953
- KATZ, Mrs. Harriet**
516 East Maryland Avenue
Phoenix, Arizona A 1950
- KAUFMAN, Bette**
9 Roseld Court
Deal, N.J. S.A. 1961
A 1961
- KAUFMANN, Elizabeth M.**
414 W. 121st Street
New York 27, N.Y. A 1950
- KAWKEWITZ, Henry (Ph.D.)**
1060 Union Street
Brooklyn 25, N.Y. A 1955
- KAY, Mrs. Victor**
1541 N. Edgemont St. Apt. 12
Los Angeles 27, Calif. A 1961
- KELSEY, Howard Phelps**
1252 Fourth Street
Sarasota, Fla. A 1944
- *KEMPLE, Camilla**
20 W. 86th Street
New York 24, N.Y. A 1940
F 1946
- KENDIG, Isabelle V. (Ph.D.)**
A 1944
Sandy Spring, Md. F 1946
- KESSLER, Mabel G. (Ph.D.)**
Super. of Special Ed.
R.D. 1
Reinholds, Pa. A 1952
- KEW, Clifton E.**
333 E. 30th St.
New York 16, N.Y. A 1949
- KIDORE, Irwin W. (Ph.D.)**
A-16 Elmwood Gardens
Coatsville, Pa. A 1959
- KING, Francis W. (Ph.D.)**
4 Kingsford Road
Hanover, N. Hampshire A 1952
- KISSINGER, R. David**
86 Dodgingtown Rd.
Bethel, Conn. St. Aff. 1960
- KITAY, Philip M. (Ph.D.)**
8707-35th Avenue
Jackson Heights 72, N.Y. A 1955
- KLASS, Walter K. (Ph.D.)**
North Central College
146 N. Sleigh Street
Naperville, Ill. A 1946
- KLATSKIN, Ethelyn H. (Ph.D.)**
Dept. of Pediatrics
4084 LMP
393 Cedar St.
New Haven 11, Conn. A 1946
F 1955
- KLEBAN, Morton H.**
Psychology Dept.
Norristown State Hosp.
Norristown, Pa. St. Aff. 1959
- KLECKNER, James H.**
70 Radnor Road
Great Neck, N.Y. St. Aff. 1961
- KLEIN, Abraham (Ph.D.)**
433 W. 21st Street
New York 11, N.Y. A 1955
- KLEIN, Mrs. Beatrice**
Woodstock, N.Y. A 1946
- *KLEIN, Eva L. (M.D.)**
1148 Fifth Avenue
New York 28, N.Y. A 1946

- KLEIN, Louis S.
280044 Morlock
Livonia, Michigan A 1952
- KLEIN, Milton
248 East 48th St.
Brooklyn 3, N.Y. S.A. 1961
- KLEINBERG, Mrs. Rosalen K.
6606 N. 11th St.
Philadelphia 26, Pa. A 1950
- KLINGER, Marjorie Pauline
1169 Eastern Pkwy. (Ph.D.)
Louisville 17, Ky. A 1961
- *KLOPFER, Bruno (Ph.D.)
Box 2971
Carmel, Calif. F 1940
- KLOPFER, Walter G. (Ph.D.)
Dept. of Psychology
Univ. of Portland A 1946
Portland 3, Oregon F 1951
- KLURFELD, Georges (M.D.)
Kloshbachstrasse 125
Zurich 7/32, Switzerland A 1961
- KNAPP, Pearl G. (Ph.D.)
Dept. of Psychiatry
Chairs of Lebanon Hosp.
4833 Fountain Avenue
Los Angeles 29, Calif. F 1956
- KOONS, Paul B. Jr. (Ph.D.)
23 Sunnyside Dr.
Athens, Ohio A 1961
- KORDA, Mrs. Geraldine J.
1080 Prospect Blvd.
Pasadena, Calif. A 1949
- KORNER, Anneliese F. (Ph.D.)
6 Parkway Road
Brookline, Mass. A 1950
F 1953
- KORNRICH, Milton
140-50 Burden Crescent
Briarwood 35, N.Y. A 1960
- KOROT, Leonard (Ph.D.)
435 N. Bedford Dr.
Beverly Hills Calif. A 1961
- KOTKOV, Benjamin (Ph.D.)
8 Orchard Street
West Brattleboro, Vt. A 1949
- KOVNAR, Murray (Ph.D.)
Director Psychology Dept.
State Hospital
Jamestown, N. Dakota
- *KRAFFT, Mrs. Margaret R.
27 West 96th Street, Apt. 4C
New York 26, N.Y. A 1940
- KRAL, V. Adalbert (M.D.)
4145 Blueridge Crescent
Montreal, Que. Can. A 1953
- KRALL, Vita (Ph.D.)
Sr. Clin. Psych.
Child Guidance Center of
Bridgeport, Inc.
1081 Iranistan
Bridgeport, Conn. A 1961
- KRAMER, George H. (Ph.D.)
722 Morgan St.
Corpus Christi, Texas A 1961
- KRAMISH, Art A. (Dr.)
Chief, Psychology Services
Federal Reformatory
U.S. Bureau of Prisons
Chillicothe, Ohio A 1957
- KRASNER, Leonard (Ph.D.)
Dept. of Psychology
Stanford University
Stanford, Calif. A 1952
- KRASS, Alvin
Pollak Clinic
Monmouth Med. Center
Long Branch, N.J. A 1961
- KRINSKY, Leonard W. (Ph.D.)
10 Stillwater Ave.
Massapequa, N.Y. A 1961
- KRUGMAN, Dorothy C. (Ph.D.)
425 Riverside Drive
New York 25, N.Y. A 1944
- KRUGMAN, Herbert E. (Ph.D.)
425 Riverside Drive
New York 25, N.Y. A 1944
- KRUGMAN, Judith I. (Ph.D.)
P.O. Box 45
Tomkins Cove, N.Y. A 1941
- *KRUGMAN, Morris (Ph.D.)
P.O. Box 45
Tomkins Cove, N.Y. F 1940
- KUSHNER, Malcolm (Ph.D.)
5500 S.W. 12th Terrace
Miami 37, Fla. A 1956
- KUTASH, Samuel B. (Ph.D.)
5 Park Road
Maplewood, N.J. A 1950
F 1951
- L'ABATE, Luciano (Ph.D.)
Dept. of Pediatrics
Wash. Univ. School of Med.
St. Louis 10, Mo. A 1961
- LACHMANN, Frank M.
235 West 77th St.
New York 24, N.Y. A 1961
- LAKIN, Harriet A.
1959 S. Crescent Heights
Los Angeles 34, Calif. A 1950
- LAMPI, Henry M.
266 Washington Ave.
Kingston, N.Y. A 1955
- LANDIS, Bernard
276 Riverside Drive
New York 25, N.Y. St. Aff 1959
- LANDISBERG, Selma
166 East 35th St.
New York 16, N.Y. A 1950
- LASAGA, Jose I. (Dr.)
Calle 9, No. 407
Vedado, Habana, Cuba A 1957
- LASKOWITZ, David
3856 Bronx Blvd.
Bronx 67, N.Y. A 1953
- LAWRENCE, Ernest S. (Ph.D.)
240 So. La Cienega Blvd.
Beverly Hills, Calif. A 1955
- LAWRENSON, Thomas J.
11-D Yale Street
Nutley, N.J. A 1955
- LAZOVIC, David A. (Ph.D.)
University of Pittsburgh
Pittsburgh 13, Pa. F 1959
- LEBEAUX, Mrs. Thelma W.
106 Newton Avenue N.
Worcester 9, Mass. A 1944
- LEBOWITZ, Anne (Ph.D.)
12971 Galewood St. St. Aff 1956
Studio City, Calif. A 1961
- LEDER, Ruth
235 E. 69th St.
New York 21, N.Y. A 1950
- LEDWITH, Nettie H. (Ph.D.)
309 Brookside Blvd.
Pittsburgh 34, Pa. F 1952
- LEE, Dorothy B.
33-35 82nd Street
Jackson Heights 72, N.Y. A 1950
- LEHMANN, Heinz E. (M.D.)
Verdun Protestant Hospital
Box 6034
Montreal, Que., Can. A 1943
F 1951
- JEHRER, Ruth (Ph.D.)
4 Washington Square Village
New York 12, N.Y. A 1944
F 1954
- LEIDEN, Irving (Ph.D.)
750 Green Bay Rd.
Winnetka, Ill. A 1956
- LFONARD, A. T.
1067 Pine St.
Muskegon, Mich. A 1954
- LEOPOLD, Julius
79 Hauch Blvd.
Riverside, Ill. St. Aff 1955
- LEFSON, David S. (Ph.D.)
Western Psychiatric Inst. and
Clinic
University of Pittsburgh
Pittsburgh 13, Pa. A 1950
- LEISER, Irving (Dr.)
University Guidance Center
University of Miami
Coral Gables 46, Fla. A 1950
- LEISER, Gerald (Ph.D.)
695 Park Ave.
New York 21, N.Y. H.F. 1961
- LEVENSTEIN, Mrs. Phyllis
3264 Island Rd.
Wanagah, L.I., N.Y. A 1940
- LEVY, Archie
University Hospital
University of Saskatchewan
Dept. of Psychiatry
Saskatoon, Can. A 1956
- LEVI, Joseph (Ph.D.)
7 E. 81st St.
New York 28, N.Y. F 1954
- LEVINE, Abraham (Ph.D.)
Hillside Hosp.
75-59 263rd St.
Glen Oaks, N.Y. A 1952
F 1951
- LEVINE, David I. (Ph.D.)
Dept. of Psychiatry
University of Nebraska
Lincoln, Nebraska F 1960
- LEVINE, Harold A. (Ph.D.)
42 Deepdale Parkway
Roslyn Heights, L.I., N.Y. F 1959
- LEVINGER, Leah
131 Jerusalem St.
Brooklyn 1, N.Y. A 1952
- LEVINSOHN, Boris M. (Ph.D.)
39-25 47th Street
Sunnyside, L.I.C. 4. F 1956
- LEVINSON, Mrs. Toby
152 Old Yonge St.
Toronto 12, Ont., Can. A 1960
- LEVIT, Herbert I. (Ph.D.)
Dixmont State Hospital
Glenfield, Pa. A 1954
- LEVY, Joshua (Ph.D.)
Clinical Psychologist
The Central Clinic
General Hospital
Cincinnati 29, Ohio A 1960
- LEVY, Ruth Jacobs (Ph.D.)
14430 Union Avenue
Cambridge Park
San Jose, Calif. A 1948
F 1951
- LEVY, Sidney J. (Ph.D.)
7417 S. Oglesby Avenue
Chicago 49, Ill. A 1956
- LEWINSON, Peter M. (Ph.D.)
5810 N. Oakland St.
Indianapolis 20, Ind. A 1958
- Lewis, Robert T. (Ph.D.)
2220 South Third Ave.
Arcadia, Calif. A 1953
- LIEBER, Mrs. Beatrice
285 Fountain Road
Englewood, N.J. A 1953
- LIEBERMAN, Mrs. Janet Chase
1136 Fifth Ave.
New York 28, N.Y. A 1956
- LIT, Jack (Ph.D.)
1172 E. Slocom
Philadelphia, Pa. A 1956
- LITTLE, Kenneth B. (Ph.D.)
Training Branch
National Inst. of
Mental Health
Bethesda, Maryland F 1959

LIU IKUS, Dr. Stanley
221 River Rd., Apt. A-5
Bogota, N.J. A 1959

LOCKWOOD, Wallace V. (Ph.D.)
2552 Fifth Ave.
San Diego 9, Calif. A 1949

LOEHRKE, Leah M. (Ph.D.)
Veterans Administration
12227 Clifton Blvd.
Lakewood, Ohio A 1954

LOUIS, Kathleen
Bureau of Child Guidance
362 Schermerhorn St.
Brooklyn 17, N.Y. A 1949

LONGLEY, James L.
Industrial Psychology Division
The Detroit Edison Company
2000 Second Avenue
Detroit 26, Mich. A 1953

LONSTEIN, Murray (Ph.D.)
Chief Psychologist VA Hosp.
Leach Farm Road
Pittsburgh 6, Pa. A 1953

*LOPES, Jose Leme (M.D.)
Rua Martins Ferreira 75
Rio de Janeiro, Brazil F 1940

LOW, Howard (Ph.D.)
VA Outpatient Clinic
17 Court St.
Boston 8, Mass. A 1954

LUCAS, Winfred B. (Ph.D.)
1019 Gayley Ave.
Los Angeles 24, Calif. A 1951

LUNDIN, William H. (Ph.D.)
1315 N. Astor St.
Chicago 10, Ill. A 1954

LYLES, William K. (Ph.D.)
47 North Clinton Ave.
Trenton 9, N.J. A 1960

LYON, LCDR W. B.
Psychiatric Unit, MCRD
Parris Island, S.C. A 1960

McBRIDE, John L.
23740 Maude Lea Circle
Novi, Michigan A 1955

MacCASLAND, Mrs. Barbara W.
Marcy State Hospital
Marcy, N.Y. A 1957

MAGNETTE, Jules (M.D.)
Nevada State Hospital A 1955
Reno, Nevada F 1961

MAHRER, Alvin R. (Ph.D.)
1315 S. Grape St.
Denver 22, Colo. A 1955

MAKSIMCZYK, Walter (Ph.D.)
1605 Eastlake Ave.
Los Angeles 53, Calif. A 1961

MALLOY, Mrs. Helga
35 Church Hill,
Westmount, Montreal, Can. A 1943

MALM, Mrs. Mildred
11423 E. Hallwood Drive
El Monte, Calif. A 1949

*MANN, Mrs. Edna B.
215 W. 98th Street
New York 25, N.Y. F 1940

MANSON, Morse P. (Ph.D.)
10644 Wilshire Blvd.
Los Angeles 24, Calif. A 1950

MANUGE, Mrs. Robt. W. M.
5077 Victoria Ave.
Montreal 29, Que., Can. A 1956

MARGOLIS, Mrs. Muriel F.
55 Fern Drive
Fast Hills
Roslyn, L.I., N.Y. A 1949

MARIANI, Eugene L. (Ph.D.)
2309 Ciniza Drive
Gallup, N. Mex. A 1958

MARIANI, Rose (Ph.D.)
2309 Ciniza Drive
Gallup, N. Mex. A 1958

MARKER, Mrs. Beatrice W.
2131 Delancey Place
Philadelphia 5, Pa. A 1951

MARKHAM, Mrs. Sylvia
116 E. 68th Street
New York 21, N.Y. A 1954

MARQUIT, Sybil (Ph.D.)
326 N.E. 26th St.
Miami, Fla. A 1961

MARSH, Donald D. (Ed.D.)
595 E. Colorado Street
Pasadena, Calif. A 1956

MARSH, James T. (Ph.D.)
Department of Psychiatry
University of California
Medical School
Los Angeles 24, Calif. A 1955

MARTIN, Harry J., Jr. (Ph.D.)
254 Leggett Drive
Abilene, Tex. A 1957

MARX, Alfred
Children's Treatment Service
Central State Hospital
Norman, Okla. A 1956

MATHER, Elise D.
5260 W. Chicago
Detroit 4, Mich. A 1948

MATHEWS, W. Mason (Ph.D.)
Merrill Palmer School
71 Ferry Avenue, E.
Detroit 2, Mich. F 1955

MATHIAS, Rudolf (Ph.D.)
5150 Juneau Road
Madison 5, Wisc. A 1950

MAYMAN, Martin (Ph.D.)
Menninger Foundation
Topeka, Kansas F 1956

*McBRIDE, Katharine E. (Ph.D.)
Bryn Mawr College
Bryn Mawr, Pa. A 1940

McCARY, James Leslie (Ph.D.)
5101 Alameda at
Southmore
Houston, Texas A 1948
F 1956

McCULLY, Robert
Payne Whitney Clinic
525 E. 68th St.
New York 21, N.Y. A 1959

McDONALD, Franklin R.
3700 Cherrywood Ave. (Ph.D.)
Los Angeles 18, Calif. A 1952

McFARLAND, Robert L. (Ph.D.)
VA Research Hospital
333 E. Huron Street
Chicago 11, Ill. F 1956

McGREGGEEV, James (Ph.D.)
Clark County Mental Health
Center
1204 Esther St.
Vancouver, Wash. A 1960

McNEILL, Harry V. (Ph.D.)
U.S.P.H.S. A 1950
42 Broadway F 1951
New York 4, New York

McPHERSON, Marion W.
350 Cloverdale (Ph.D.)
Akron 2, Ohio A 1953

MEHR, Helen Margulies (Ph.D.)
2264 Fairhill Lane A 1941
San Jose, Calif. F 1949

MEKLER, Sara
Gutemberg 231
Mexico 5, D.F. A 1961

MENDENHALL, John H.
220 E. Church St.
Xenia, Ohio A 1958

MERCER, Margaret (Ph.D.)
St. Elizabeths Hospital A 1946
Washington, D.C. F 1950

MESSERSCHMIDT, Ramona
Mental Hygiene Clinic (Ph.D.)
VAH, Spokane, Wash. A 1961

MEYER, George (Ph.D.)
2479 16th Avenue
San Francisco, Calif. A 1950

*MEYER, Mortimer M. (Ph.D.)
503 N. Bronson Avenue A 1940
Los Angeles 4, Calif. F 1949

*MIALE, Dr. Florence R.
225 West 86th Street
New York, N.Y. F 1940

MICHAEL, Carmen Miller
6223 Lupton A 1951
Dallas 25, Texas F 1955

MICHAEL-SMITH, Harold
1230 Park Avenue (Ph.D.)
New York 28, N.Y. F 1955

MIERZWA, Mr. John
Testing & Counseling Center
University of Texas
Austin 12, Texas Aff. 1960

MILLER, Cecil R.
1271 Midvale Ave. St. Aff. 1953
Los Angeles 24, Calif. A 1955

MILLS, David H.
914L Cherry Lane
E. Lansing, Mich. A 1960

MILSTEIN, Dr. A. Freda
17558 Prest
Detroit 35, Mich. A 1946

MIMS, Mrs. Jean Giesey
1110 E. 32nd Street
Austin, Texas A 1943

MINDESS, Harvey (Ph.D.)
2108 Balsam
Los Angeles 25, Calif. A 1953

MINDLIN, Mrs. Dorothea F.
6408 Bannockburn Dr.
Bethesda 14, Md. A 1955

MOCK, Joe F. (Ph.D.)
615 Sixth Ave.
Huntington, W. Va. A 1960

MOLISH, Dr. Herman B.
Medical Res. Lab. Box 100
U. S. Sub Base
New London, Conn. A 1950

MOORE, Mrs. Harriet Bruce
145 E. Ohio Street
Chicago 11, Ill. A 1953

MORELAND, Mrs. Margaret E.
7 Burkewood Road
Hartsdale, N.Y. A 1950

MORF, Gustav (M.D.) (Ph.D.)
1410 Faville Avenue
Montreal 19, Quebec
Canada A 1954

MORGAN, Clellen L. (Ph.D.)
R.D. No. 20, Beatty Road
Media, Pa. A 1950

MORGAN, Olive J. (Ph.D.)
R.D. No. 20, Beatty Road
Media, Pa. A 1950

MORIZE, Mrs. Andre
4, Rue Jean-du-Bellay
Paris 4^e, France A 1945

MORRIS, Charles M. (Ph.D.)
73 North Second St.
Easton, Pa. A 1956

MORRISON, Alfonso
Child Care Center
2110 Fifth St. St. Aff. 1956
Golena Park, Texas

MORROW, J. Lloyd (M.D.)
197 Passaic Avenue
Passaic, N. J. A 1943

MOTOAKI, Dr. Hiroshi
3-10 Honochodori
Nakano-Ku-Tokyo F 1961

MOTZ, Dr. Gerald
1604 Holbrook
Ponca City, Oklahoma A 1954

MUELLER, Adolph R. (M.D.)
1220 Washington Street
Leavenworth, Kans. A 1943

- MUFNCH, George (Ph.D.)
San Jose State College
Bldg K, Room 1a
San Jose, Calif. A 1946
- MULLEN, Esther (Ph.D.)
10 Downing Street
New York 14, N. Y. A 1950
- *MUNROE, Ruth L. (Ph.D.)
289 Central Park West
New York 24, N. Y. F 1940
- MUNZ, Alan
6536 99th St.
Forest Hills, L. I., N. Y. A 1955
- MURPHY, Rev. Kenneth
26 S. Center Street
Springfield, Ohio Aff. 1954
- MURPHY, Lois Barclay (Ph.D.)
Menninger Foundation
Topeka, Kans. A 1941
- MURRAY, David (Dr.)
VA Mental Hygiene Clinic
Chimes Bldg.
500 So. Salina St.
Syracuse 2, N. Y. A 1958
- MURRAY, Henry (M.D.)
7 Divinity Avenue
Cambridge 38, Mass. F 1950
- MURSTEIN, Dr. Bernard I.
Interfaith Counseling Center
729 Southwest Alder St. A 1957
Portland 5, Ore. F 1959
- MUSIKER, Harold (Ph.D.)
307 Wayland Ave.
Providence 6, R.I. A 1961
- NAPOLI, Peter J. (Ed.D.)
West Shore Rd., RFD 2
Putnam Valley, N.Y. A 1949
- NELSON, Mrs. Martha
107 E. Spring St. E.
Oxford, Ohio A 1961
- NEUMAN, Gerard G. (Ph.D.)
2925 Arrowwood Trail
Deerfield, Illinois A 1955
- NEWMAN, Joseph (Ph.D.)
University Drive
Pittsburgh 40, Pa. A 1950
- NICHOLAS, Alma L.
200 Retreat Avenue
Hartford 2, Conn. A 1954
- NOBLE, Eric (Ph.D.)
702 First Ave. So.
Fargo, N. Dakota A 1961
- NORTHCOTT, Hollie
1452 Willard St.
San Francisco 17, Calif. A 1954
- NOSAL, Walter S. (Ed.D.)
John Carroll University
Cleveland 18, Ohio A 1954
- NUNFZ, Rafael (Ph.D.)
Facultad de Filosofia
y Letras, Univ. Nacional
Autonoma de Mexico
Ciudad Universitaria A 1954
Mexico 20, D.F., Mexico F 1958
- NUTT, Frances D.
Box 938
Eloy, Ariz. A 1958
- OCHROCH, Ruth (Dr.)
50 E. 90th St.
New York 28, N.Y. A 1950
- ODOM, Charles L. (Ph.D.)
404 Carondelet Bldg.
New Orleans 12, La. A 1949
- OLIN, Tom Davis
Inst. of Living
200 Retreat Ave.
Hartford 2, Conn. A 1956
- OLINGER, Leonard Bennett
416 No. Bedford Dr. (Ph.D.)
Beverly Hills, Calif. A 1954
- O'REILLY, P. Oliver,
(M.B., B.Ch.)
Dir., Psychiatric Dept.
Union Hospital, 3rd Floor
Moose Jaw
Saskatchewan, Can. A 1957
- ORCEL, Sidney A. (Ph.D.)
205 Crawford Ave.
Syracuse 3, New York A 1952
- ORR, David Hamilton (Ph.D.)
Psychological Services
Allentown State Hospital
Allentown, Pa. F 1956
- OSSORIO, Abel Garcia (Ph.D.)
Department of Psychology
Washington University
St. Louis, Mo. A 1951
- OSTERWEIL, Jerry (Ph.D.)
Student Health Service
Univ. of Penna. Hospital
Philadelphia 4, Pa. F 1960
- PAINTING, Donald H. (Ph.D.)
1936 Cardinal Ave.
West Chester, Pa. A 1960
- PAIM, Rose (Ph.D.)
263 West End Avenue
New York 23, N. Y. F 1955
- PAOLINO, Albert (Dr.)
12004 Paul Ave.
Cleveland 6, Ohio A 1958
- PARKER, Rolland S.
30 A Joralemon St.
Brooklyn 2, New York A 1957
- PARNICKY, Dr. Joseph J.
Box 20
Bordentown, N.J. A 1949
- PAUL, Waters C.
Psychology Dept. Boyce Hosp.
Station 3
Tuscaloosa, Ala. A 1956
- *PAULSEN, Alma A. (Ph.D.)
667 Madison Ave.
New York 21, N.Y. F 1940
- PAYNE, David H., M.A.D.O.
8345 Talbert Avenue
Huntington Beach, Calif. A 1950
- PEAK, Horace M.
549 N. Arrowhead
San Bernardino, Calif. A 1949
- PEARSE, Robert F.
Tiffany Road
Norwell, Mass. A 1961
- PECK, Miss Rosalind
317 F. 24th St.
New York 10, N.Y. A 1958
- PECKARSKY, Adeline (Ph.D.)
67 Parker Ave.
Maplewood, N.J. A 1958
- PEIXOTTO, Helen E. (Ph.D.)
Child Center
Catholic University of America
Washington, D.C. F 1955
- PELLAT, Marjorie L.
503 C Avondale Ave.
Los Angeles 49, Calif. A 1961
- PENA, Cesareo D. (Ph.D.)
Mental Health Institute
Box 111
Independence, Iowa A 1951
- PESETSKY, Fred J.
Box 902
Norfolk, Nebr. St. Aff. 1957
A 1961
- PETERS, Marie Wilson
6144 Wayne Avenue
Philadelphia 44, Pa. A 1947
F 1951
- PEYMAN, Douglas A. R.
Psychology Department (Ph.D.)
Alabama State Hospital
Tuscaloosa, Ala. F 1956
- PHILIP, Anthony F. (Ph.D.)
Austen Riggs Center
Stockbridge, Mass. A 1960
- PHILLIPS, John (Ph.D.)
4 Andrews Road
Malvern, Pa. A 1959
- PHILLIPS, Leslie (Ph.D.)
Worcester State Hospital
Worcester 1, Mass. H.F. 1961
- PHILLIPS, Maurice
92 a Nassau St.
Princeton, N.J. A 1961
- PHILLIPSON, Mr. H.
Tavistock Clinic
2 Beaumont St.
London W 1, England H.F.
1961
- *PIOTROWSKI, Zygmunt A.
Jefferson Medical College
1025 Walnut St.
Philadelphia 7, Pa. F 1940
- PLATT, Henry (Ph.D.)
The Devereux Foundation
Inst. for Res. and
Training
Devon, Pa. A 1950
F 1961
- PLITTMAN, Jack C.
Psychology Department
Patton State Hospital
San Bernardino, Calif. A 1951
- POPPELSTONE, John (Ph.D.)
350 Gloverdale
Akron 2, Ohio A 1958
- PORTER, Mrs. Lucille S.
250 First Avenue
New York 9, N. Y. St. Aff. 1953
- POSER, Ernest George (Ph.D.)
Department of Psychology
McGill University
3684 Mt. Tavish Street
Montreal, P.Q., Canada A 1950
F 1953
- POTTHARST, Karl E. (Ph.D.)
4911 Van Nuys Blvd.
Suite 506
Sherman Oaks, Calif. F 1957
- POWELL, David (Ph.D.)
71 West 12th St.
New York 11, N.Y. A 1961
- PRENTICE, Norman M.
Judge Baker Guidance Center
295 Longwood Ave.
Boston, Mass. A 1961
- PRICE, Mrs. Marian Blewitt
854 S. Euclid
Pasadena 5, Calif. A 1953
- PROCTOR, Paul W. (Ph.D.)
44 West 10th St.
New York 11, N.Y. A 1951
- PUZZO, Frank S.
130-52 232nd Street
Laurelton, N. Y. A 1949
- RAAUM (Ph.D.)
Acting Chief Psych.
Dammsch State Hosp.
Wilsonville, Ore. A 1959
- *RABIN, Albert I. (Ph.D.)
Department of Psychology
Michigan State University
East Lansing, Mich. F 1955
- RADER, Dr. Gordon E.
Psychology Section
Department of Psychiatry
North Carolina
Memorial Hospital
Chapel Hill, N.C. A 1956
- RADTKE, William L.
Manatee Junior College
Bradenton, Florida A 1954
- RAIFMAN, Irving (Ph.D.)
3102 Woodhollow Drive
Chevy Chase, Maryland
- RAINWATER, Mrs. H. B.
Moore
Newmans Park 19
Hamburg-Nienstedten,
Germany A 1956

- RAINWATER, Lee (Ph.D.)**
145 E. Ohio St.
Chicago, Ill. A 1956
- RAPKIN, Maurice (Ph.D.)**
800 S. Robe, 1909 Bldg
Los Angeles 35, Calif. A 1952
- RAPPAPORT, Sheldon R. (Ph.D.)**
290 Linden Lane
Merion, Pa. A 1951
F 1956
- RAPPAPORT, Sidney M. (Ph.D.)**
1423 Mellon Road
Wyncote, Pa. A 1949
F 1953
- RAUB, Edwin S. (Ed.D.)**
213 Midland Avenue
Wayne, Pa. A 1952
- RECORD, Father Maurice A.**
Dept. of Psychology
Assumption College of
Windsor
Windsor, Ontario, Can. A 1954
- REED, Max (Ph.D.)**
University of Portland
Portland 3, Ore. F 1961
- REES, Mrs. Renee G.**
27 West 96th St.
New York 25, N.Y. A 1959
- REICHARD, Suzanne (Ph.D.)**
1075 Cragmont Avenue
Berkeley, Calif. A 1941
- REICHENBERG-HACKETT, Wally (Ph.D.)**
c/o Department of Psychology
College Station
Duke University
Durham, N. C. A 1948
- REIS, Walter J. (Ph.D.) (M.D.)**
552 Neville
Pittsburgh, Pa. A 1943
- REISEL, Jerome (Ph.D.)**
4937 Nagle Avenue
Sherman Oaks, Calif. A 1955
- REISS, William J. (Ph.D.)**
VA Center
Kecoughtan, Va. A 1955
- REITZ, Mrs. Edna Maisner**
2314 Clarendon
Woodland Hills, Calif. A 1953
- REITZELL, Mrs. Jeanne M.**
500 S. Arroyo Boulevard
Pasadena 2, Calif. A 1949
- RESWICK, Mrs. Serena W.**
86 Saddle River Rd.
Monsey, N.Y. A 1950
- RHODES, Mrs. Charlotte**
70 La Salle St.
New York 27, N.Y. A 1961
- RICHARDS, T. W. (Ph.D.)**
Kennedy Child Study Center
2024 Arizona St.
Santa Monica, Calif. F 1954
- RICKERS OVSIANKINA, M.**
Univ. of Connecticut
Storrs, Conn. F 1940
- RIDGE, Bradley B.**
Chatham Court
49 St. & Locust
Philadelphia 39, Pa. A 1958
- RISCH, Frank (Ph.D.)**
3097 Manning Avenue
Los Angeles 64, Calif. A 1949
- *RITEY, Hector J. (M.D.)**
815 Park Avenue
New York 21, N.Y. A 1940
- RITZ, George H. Jr. (Ph.D.)**
4070 Mayfield Road
Cleveland 21, Ohio A 1959
- *RIVERS, Mrs. Hubert M.**
111 Mitchell Drive
Pittsburgh 28, Pa. A 1940
- ROACHE, Miriam Haines**
2329 Hudson Terrace (Ph.D.)
Covetsville, Ft. Lee, N.J. A 1951
- ROBINOWITZ, Ralph (Ph.D.)**
Psychology Service
VA Hospital
Dallas 16, Texas A 1957
- ROBINSON, Elizabeth Foster**
Child Psychiatry Division
University Hospitals
Madison, Wis. A 1954
- ROCKBERGER, Harry (Ph.D.)**
299 So. Harrison St.
East Orange, N. J. A 1954
- RODAN, Mrs. Henrietta Itta**
510 E. 85th Street
New York 28, N.Y. St. Aff. 1954
- RODRIGUEZ-VALDERRAMA, Dr. Jose**
Calle 71-#10-53
Bogota, D.E., Colombia
- ROE, Anne (Ph.D.)**
Graduate School of Education
Harvard University
Cambridge 38, Mass. F 1959
- ROGERS, Lawrence S. (Ph.D.)**
1046 Madison Street
Denver 6, Colo. F 1954
- ROLAND, Alan O. (Ph.D.)**
71 West 12th St.
New York 14, N.Y. A 1960
- ROOK, Leroy H.**
205 N.E. 8th
Fulton, Mo. A 1957
- RORSCHACH, Mme. Olga**
Hirschgartenweg
Zurich 57, Switzerland H. M. 1954
- ROSE, Margaret E.**
34 St. Marks Place
New York 3, N.Y.
- ROSE, Nicholas**
921 Westwood Blvd.
Los Angeles, Calif. A 1956
- ROSEN, Ephraim (Ph.D.)**
Dept. of Psychology
University of Minnesota
Minneapolis 14, Minn. H.F. 1961
- ROSEN, Esther Katz (Ph.D.)**
1810 Rittenhouse Sq.
Philadelphia 3, Pa. A 1945
F 1951
- ROSENBERG, Israel H.**
89-38 Whitney Avenue
Elmhurst 73, L. I., N. Y. A 1953
- ROSENTHAL, Robert (Ph.D.)**
Psychology Dept.
Univ. of No. Dakota
Grand Falls, N.D. A 1957
F 1959
- ROSENTHAL, Vin (Ph.D.)**
Dept. of Neurology
and Psychiatry
Northwestern University
Medical School
301 E. Chicago Avenue
Chicago 11, Ill. A 1956
- ROSNFR, Stanley (Ph.D.)**
971 Fairfield Avenue
Bridgeport, Conn. A 1954
- ROSS, Alan O. (Ph.D.)**
Pittsburgh C.G.C.
201 De Soto St.
Pittsburgh 13, Pa. F 1960
- ROSS, M. Eleanor (Ph.D.)**
B-910 Presidential Apts. A 1944
Philadelphia 31, Pa. F 1949
- *ROSS, W. Donald (M.D.)**
Department of Psychiatry
Cincinnati General Hospital
Cincinnati 29, Ohio F 1940
- ROTHENBERG, Mrs. Eleanore**
77 West 85th St.
New York 24, N.Y. A 1958
- ROTMAN, Saul R. (Ph.D.)**
28 Allendale Ave.
St. Cloud, Minn. A 1947
- RUHL, Mrs. R. Ernest**
817 Chestnut Street
Millinburg, Pa. A 1947
- RUIZ, Rene Arthur**
2721 P. St.
Lincoln, Neb. A 1961
- RUJA, David H. (Ph.D.)**
435 No. Roxbury Drive
Beverly Hills, Calif. A 1949
- RUSSELL, Howard**
3680 Fairway Blvd.
Los Angeles 43, Calif. A 1954
- *RYMER, Charles A. (M.D.)**
230 Majestic Bldg.
Denver 2, Colo. F 1940
- SABATH, Gerald**
9 Patchin Pl.
New York 11, N.Y. A 1960
- ST. CLAIR, Walter F. (Ed.D.)**
999 Mammoth Road
Manchester, N. H. A 1943
- SALTZMAN, Marguerite R.**
2037 Spruce Street
Philadelphia 3, Pa. A 1950
- SALTZMAN, Sara**
7012 Wilson Lane
Bethesda 14, Md. A 1950
- SALZMAN, Mrs. Anne**
12548 Everglade Street
Los Angeles 66, Calif. A 1953
- SANCHEZ-GARCIA, Dr. Jose**
Casilla de Correos 57
Miraflores, Lima, Peru A 1959
- SANDER, Emilie T.**
657 W. 161st Street
New York, N. Y. A 1950
- SANDERS, David H.**
1032 Remington Dr.
Bldg. 8, Apt. 2
Sunnyvale, Calif. St. Aff. 1960
- SANDERSON, Herbert (Ph.D.)**
1825 34th Avenue
San Francisco 22, Calif. A 1952
- SANFORD, Nevitt (Ph.D.)**
Vassar College
Poughkeepsie, N.Y. F 1959
- SARGENT, S. Stansfeld (Ph.D.)**
4510 E. Vermont Ave.
Phoenix, Ariz. A 1961
- SCALES, Margaret B. (Ph.D.)**
50 King St.
New York 14, N.Y. A 1955
- SCHACHT, Mrs. Leatrice Styrt**
5 Avis Drive
New Rochelle, N. Y. A 1950
- SCHACHTTEL, Ernest G.**
299 Riverside Drive
New York 25, N. Y. F 1951
- SCHACHTTEL, Mrs. Zeborah**
299 Riverside Drive
New York 25, N. Y. A 1953
- SCHACHTER, M. (M.D.)**
40 A Boulevard Voltaire
Marselle 1, France H.F. 1961
- SCHAFER, Robert E. (Ph.D.)**
Harless & Kirkpatrick
Associates
420 W. Lafayette Street
Tampa, Fla. A 1956
- SCHANBERGER, William J.**
305 Veri Avenue
Pittsburgh 20, Pa. A 1944
- SCHATTMAN, Mrs. Esther**
Preger
210 E. 68th Street
New York 21, N. Y. A 1950

- SCHAW, Louis G.
University of Illinois
College of Med. Psychiatry
912 S. Wood St.
Chicago 12, Ill. A 1959
- SCHIR, Sam C. (Ph.D.)
1668 Portland Avenue
St. Paul 4, Minn. A 1956
- SCHILLINGER, Morton
441 West End Avenue
New York 24, N. Y. A 1953
- SCHLESINGER, Mrs. Alicia de
Solis 155, VIII/A A 1940
Buenos Aires, Argentina
- SCHMALZRIED, Newell T.
(Ph.D.)
412 S. Boots St.
Marion, Ind. A 1956
- SCHMIDT, Fritz (M.S., Dr. Jur.)
6056 Upland Terrace A 1942
Seattle 18, Wash. F 1945
- SCHNEIDER, Stanley F. (Ph.D.)
Neuropsychiatric Institute
University Hospital
Ann Arbor, Mich. A 1954
- SCHOENHORN, Miss Elizabeth
Reinisch 4
Vienna, Austria
- SCHON, Martha (Ph.D.)
10 West 86th St.
New York 24, N. Y. F 1961
- SCHONBAR, Rosalea Ann
(Ph.D.)
30 W. 60th St., Apt. 10H
New York 23, N. Y. A 1943
- SCHUBERT, Herman J. P.
(Ph.D.)
Route 2
500 Klein Road
Buffalo 21, N. Y. A 1950
- SCHULMAN, Irving (Ph.D.)
638 Harvard Road
Cynwyd, Pa. A 1952
- SCHUMACHER, Audrey Sims
2257 Northwest 11th Ave.
Gainesville, Fla. F 1949
- SCHUMACHER, Henry C.
(M.D.)
2257 Northwest 11th Ave.
Gainesville, Fla. A 1941
- SCHWARTZ, Arthur A.
251 Central Park West
New York 24, N. Y. A 1951
- SCHWARTZ, Emanuel K.
12 E. 87th St. (Ph.D., D.S. Sc.)
New York 28 A 1949
New York F 1952
- SCHWARTZ, Mrs. Lita
411 Lodges Lane
Elkins Park 17, Pa.
St. Aff. 1960
- SCHWERIN, Mrs. Erna
1102 W. Elm St.
Lima, Ohio A 1950
- SCOTT, Edward M. (Ph.D.)
3632 N.E. Davis
Portland 12, Oregon
- SEIDENFELD, Morton A. (Ph.D.)
5410 Connecticut Ave., N.W.
A 1944
Washington 15, D.C. F 1954
- SEILER, Mrs. Geraldine F.
1120 Oakland Drive A 1946
Mount Dora, Florida F 1950
- SEITZMAN, Daniel
2387 Ocean Avenue
Brooklyn 29, N. Y. A 1949
- SELIG, Kalman (Ph.D.)
22 Ball Street
Irvington 11, N. J. A 1950
- SELZER, Samuel (Ph.D.)
Box 65
Sonyea, N. Y. A 1954
- SELZER, Samuel (Ph.D.)
Minnesota Mental Health
Center
Albert Lea, Minn. A 1956
- SHACKETT, Mrs. Sarah Eyre
Route 2, Box 35
Espanola, New Mexico A 1942
- SHAH, Saleem A. (Ph.D.)
1030 Kent Avenue
Catonsville 28, Maryland A 1958
- SHAPIRO, David (Ph.D.)
Austen Riggs Foundation
Stockbridge, Mass. A 1950
- SHARPE, Susie McMillan (Ph.D.)
46 W. 4th Street
Mt. Vernon, N. Y. A 1948
- SHEEHAN, Joseph (Ph.D.)
416 - 21 Place
Santa Monica, Calif. A 1952
- SHERMAN, Murray H. (Ph.D.)
350 Central Park West
New York 25, N. Y. A 1956
- SHIPMAN, William G. (Ph.D.)
3601 Fifth Avenue
Pittsburgh 13, Pa. A 1956
- SHNEIDMAN, Edwin S. (Ph.D.)
11451 Kingsland Avenue
A 1949
Los Angeles 66, Calif. F 1951
- SHOR, Joel (Ph.D.)
635 Castle Hill Ave. 4B
Bronx 72, N. Y. A 1945
- SHULMAN, Harold S. (Ph.D.)
Mental Health Clinic
501 E. Springfield Ave.
Champaign, Ill. A 1957
- SIEGEL, Joseph H. (Ph.D.)
11330 Hillcrest Road
Dallas 30, Texas A 1956
- SIEGEL, Max (Ph.D.)
50 Kenilworth Place
Brooklyn 10, N. Y. F 1956
- SIEGEL, Miriam G. (Ph.D.)
57 E. 90th Street
New York 28, N. Y. F 1949
- SILVERMAN, Lloyd (Ph.D.)
680 West End Avenue
New York 25, N. Y. A 1960
- SILVERSTEIN, Mrs. Sophie M.
2301 Kings Highway
Brooklyn 29, N. Y. St. Aff. 1956
- SIMKIN, James S. (Ph.D.)
435 No. Bedford Drive
Suite 102
Beverly Hills, Calif. A 1952
- SIMKINS, Lawrence
Florida State Univ.
Department of Psychology
Tallahassee, Fla. St. Aff. 1958
- SIMON, Maria D. (Ph.D.)
Hartackerstrasse 44
Vienna 19, Austria A 1959
- SINGER, Erwin (Ph.D.)
53 Dunham Road
Hartsdale, New York F 1959
- SINGER, Roland H.
7895 Mark Dr.
Verona, Pa. A 1953
- SISSON, Boyd D. (Ph.D.)
2237 Darlington Dr.
Forest Acres,
Augusta, Ga. F 1957
- SKEELS, Dell
Humanistic-Social Department
University of Washington
Seattle 15, Wash. Aff. 1954
- SLESS, Bernard
225 Upland Road
Merion, Pa. A 1952
- SMITH, Frances (Ph.D.)
7155 E. Wardlow Road
Long Beach 8, Calif. A 1955
- SMITH, Mrs. Margaret J.
Wisconsin Diagnostic Center
1552 University Avenue
Madison, Wis. A 1950
- SMOLINSKY, Harold J. (Ph.D.)
Wernersville State Hospital
Wernersville, Pa. A 1952
- SNOWDEN, Robert F.
2540 Huntington, Dr.
San Marino, Calif. A 1953
- SOBOL, Albert L. (Ph.D.)
308 Betsy Brown Road A 1949
Port Chester, N. Y. F 1954
- SOLI, Jerome
33-45 90th St.
Jackson Hts. 72, N. Y. SA 1961
- SOMERVILLE, Addison W.
5439 So. Greenwood Ave.
Chicago 15, Illinois A 1956
- SOMMERS, Vita Stein (Ph.D.)
931 So. Windsor Blvd.
Los Angeles 19, Calif. A 1946
- SOSNOFF, Mrs. Miriam
372 Central Park West
New York 25, N. Y. A 1955
- SPANAY, Emma (Ph.D.)
Queens College
Flushing 67, N. Y. A 1949
- SPENCER, Mrs. Betty L.
1912 18th Street
Huntington, W. Va. A 1951
- SPIEGELMAN, J. Marvin (Ph.D.)
420 North Camden Drive
Beverly Hills, Calif. A 1953
- SPIN, Mrs. Lillian
500 E. 56th Street
Brooklyn 5, N. Y. A 1950
- SPINDLER, Mrs. Joan Elizabeth
201 Kindersley Ave.
Town of Mount Royal
Quebec, Canada A 1948
- SPIER, Dr. Jess
Box 8186, Univ. Guidance
Center
Univ. of Miami
Coral Gables, Fla. F 1958
- SPIRES, Alan M. (Ph.D.)
Provincial Hospital
Lancaster, N.B., Can. A 1954
- SPRINGER, Florence E.
370 W. 255th Street
Bronx 71, N. Y.
- STANFORD, Dr. Margaret J.
Sonoma State Hospital
Eldridge, Calif. A 1950
- STANTON, Mrs. Harriet
15 Livermore Road A 1942
Wellesley Hills 82, Mass.
- STAVRIANOS, Mrs. Bertha
823 Ingleside Place
Evanston, Ill. A 1943
- STEEN, Thomas W. (Ph.D.)
11153 Rosarita Drive
Loma Linda, Calif. A 1954
- STEIN, Morris I. (Ph.D.)
21 Washington Pl.
New York 3, N. Y. F 1959
- STEINER, M. Elizabeth
220 Brookdale Avenue
Newark 6, N. J. F 1946
- STEINER, Meta (Ph.D.)
40-70 Hampton Street
Elmhurst 73, N. Y. F 1950
- STEINZOR, Bernard (Ph.D.)
365 West End Ave.
New York 24, N. Y. A 1943

- STENDEL, Mrs. Kathleen**
Glendon-Westwood Med. Bldg.
1250 Glendon Avenue
Los Angeles 24, Calif. A 1950
- STERN, Mrs. Kathryn Werner**
14-32 30th Drive
Astoria 2, L.I., N.Y. A 1952
- STERNBERG, David**
Bureau of Mental Health
235 West 23rd St.
New York, N.Y. A 1961
- STERNE, Spencer B.**
9157 Condesa Dr.
Sacramento 26, Calif. A 1953
- STIRLING, Miss Elizabeth**
2024 R. Street, N.W.
Washington 9, D.C. St. Aff 1960
- STONE, Irving R.**
State Mental Hygiene Clinic
3525 Fourth Avenue
San Diego 3, Calif. A 1951
- *STONE, L. Joseph (Ph.D.)**
Vassar College
Poughkeepsie, N. Y. F 1961
- STOOPS, Mrs. Wanda Rah**
4256 Knollton Rd.
Indianapolis 5, Ind. A 1949
- STOTZ, Marion**
340 W. Enid Drive
Key Biscayne 49, Fla. A 1953
- STRAIT, Bennett**
509 Scott Street
Stroudsburg, Pa. A 1950
- STRAUSS, Mrs. Elsa L.**
3819 Dakota Street
Cincinnati, Ohio A 1951
- STRUTHERS, Alice Ball (Ph.D.)**
2501 Palos Verdes Drive, N.
Palos Verdes Estates
California A 1949
- STURCH, Jack E. (Ph.D.)**
Inst. of Psychological Services
Illinois Institute of Technology
Chicago 16, Ill. A 1961
- STURGEON, Artie (Ph.D.)**
St. Albans Hospital
Radford, Va. A 1961
- SULZER, Edward S.**
Box 393, Univ. of Minnesota
Medical School St. Aff. 1954
Minneapolis 14, Minn. A 1958
- SUNDBERG, Norman D.**
Dept. of Psychology
University of Oregon
Eugene, Oregon F 1961
- SWIFT, Joan Woodcock (Ph.D.)**
5628 S. Blackstone Avenue
Chicago 37, Ill. A 1945
- TABIN, J. K.**
162 Park Ave.
Glencoe, Ill. A 1952
- TADA, Mr. Haruo**
2-26 Ikeda-machi
Kanazawa-shi
Ishikawa-ken, Japan
- TALLEN, Norman (Ph.D.)**
VA Hospital
Northampton, Mass. A 1953
- TANAKA, Mr. Fujio**
3-7 Ishizakakakuba
Kanazawa, Shi
Ishikawa-ken, Japan A 1959
- TAULBEE, Earl S. (Ph.D.)**
Veterans Adm. Hospital A 1953
Lincoln, Nebraska F 1955
- TAYAL, Shanti**
2025 Eye St., N.W.
Apt. 1025
Washington 6, D.C. A 1960
- TAYLOR, Mrs. Verda**
1307 Maple Ave.
Lancaster, Pa. A 1950
- *TEICH, Mrs. Marianne**
148-45 89th Ave. Apt. C4
Jamaica 35, N.Y. A 1940
- TEICHER, Arthur, Dr.**
215 West 88th St.
New York 24, N.Y.
- TEMERLIN, Maurice K.**
Guidance Service (Ph.D.)
University of Oklahoma
Norman, Okla. A 1956
- TENNEY, Edward Vernon**
(Ph.D.)
735 East Holland Ave.
Fresno 4, Calif. A 1948
- THETFORD, William N.**
(Ph.D.)
9 E. 78th St.
New York 21, N.Y. F 1957
- THORNTON, Thomas E.**
(Ph.D.)
Route 1, Box 223
Waukegan, Illinois A 1959
- TOLMACH, Mrs. Regina E.**
16 W. 77th Street
New York 24, N. Y. A 1949
- TOMBLEN, Donald (Ph.D.)**
44 South Munn Ave.
East Orange, N.J. A 1956
- TOMKINS, Silvan S. (M.D.)**
Princeton University
36 College Road
Princeton, New Jersey
- TOPPING, Marion Powers**
(Mrs. Robert C.)
122 East 22nd Street
New York 10, N.Y. A 1948
- TOWNSEND, Mrs. Marjorie M.**
Plainfield, Vt. A 1949
- TRACHTMAN, Gilbert M.**
(Ph.D.)
2941 Carlyle Road
Wantagh, L.I., N.Y. A 1954
- TREAT, Wolcott C. (Ph.D.)**
5027 Campanile Drive
San Diego 15, Calif. A 1953
- TRIPP, Clarence A. (Dr.)**
The Handwriting Inst.
18 E. 48th St.
New York 17, N.Y. A 1956
- TROUP, Evelyn (Ph.D.)**
12890 Sunset Blvd.
Los Angeles 49, Calif. F 1949
- *TUFT, Carolyn M. (Ph.D.)**
4613 Larchwood Avenue
Philadelphnia, Pa. A 1940
- TYRREL, Mrs. Marcel J.**
Director of Psych. Services
Boys' Training Center
So. Portland, Me. A 1960
- ULLMANN, Dr. Leonard P.**
VA Hospital
Palo Alto, Calif. A 1958
- UMPIERRE, Francisco Jose, Dr.**
Independencia 565 St. Aff.
Baldrich Hato Rey, P.R. 1956
- VACCARO, J. John (Ph.D.)**
5 Dartmouth St.
Forest Hills 75, N.Y. F 1955
- VALERIUS, Elizabeth**
285 Blecker St., Apt. 3
New York 14, N.Y. S.A. 1961
- VANDENBERG, Steven G.**
(Ph. D.)
Dept. of Pediatrics
Univ. of Louisville
323 Chestnut St.
Louisville, Ky. A 1951
- VAN WEST, Mrs. Joan**
424 E. 14th St., Apt. 3 C
New York 9, N.Y. St. Aff. 1954
- VASSILIOU, Mrs. Vasso**
911 West Carmen St.
Chicago 40, Ill. A 1960
- VAYHINGER, Dr. John M.**
Dept. of Pastoral Psych.
and Counseling
Garrett Biblical Inst.
Northwestern Univ. Campus
Evanston Illinois A 1952
- VERRILL, Bernard (Ph.D.)**
Department of Psychology
Bradley University
Peoria, Ill. A 1958
- VINSON, David (Ph.D.)**
1210 Medical Towers
Houston 25, Texas A 1949
- VOGEL, Horst (Dr.)**
Institute for Psychoanalysis
Biebergrasse 8, Germany A 1957
- VORHAUS, David**
27 W. 86th Street
New York 24, N. Y. H. M. 1954
- VORHAUS, Pauline G. (Ed.D.)**
27 W. 86th Street
New York 24, N. Y. F 1944
- WAGNER, Edwin**
2641 Chamberlain Ave.
Apt. 16
Akron 13, Ohio A 1961
- WAGNER, Mazie Earle (Ph.D.)**
Route 2
500 Klein Road
Buffalo 21, N.Y. A 1950
- WAGNER, Nathaniel N. (Ph.D.)**
Astor Home for Children
Rhinebeck, N.Y. A 1961
- WALD, Charles**
23 Bond St.
Great Neck, N.Y. A 1961
- WALLER, Patricia (Ph.D.)**
Dept. of Psychology
VA Hospital
Brockton, Mass. A 1961
- WALTON, Mrs. Norma R.**
930 Dart Road, Rt. 3
Mason, Michigan A 1949
- WARNER, Samuel J. (Ph.D.)**
200 East 16th Street
New York 3, N.Y. A 1953
- WARRFN, Lurene Z.**
Mt. Pleasant State Home &
Training School
Mt. Pleasant, Mich. A 1949
- WARSHAWSKY, Mrs. Florence**
2889 Torrington Road A 1949
Shaker Heights 22, Ohio
- WATERS, Thomas J.**
812 S. Edgewood St. Aff. 1955
Columbia, Mo. A 1961
- WATKINS, Roberta Frank**
533 San Marino
San Marino, Calif. St. Aff. 1955
- WEINSTEIN, Dr. Marvin S.**
405 Nova Albion Way
San Rafael, Calif. A 1958
- WEINTRAUB, Irwin G.**
1300 Delaware Ave.
Wilmington 6, Del. A 1961
- WEIR, John R. (Ph.D.)**
2841 Highview Avenue
Altadena, Calif. A 1954
- WEISS, Bertram A. (Ph.D.)**
714 Cedar
Hastings, Neb. A 1956
- WEISS, Eralyn R.**
1020 Centre Ave.
Reading, Pa. A 1950
- WEISS, Herman R. (Ph.D.)**
1277 E. 48th Street
Brooklyn 34, N. Y. A 1953
- WEISS, Sheldon W. (Ph.D.)**
604 Clearview Ave.
Wycliff
Wilmington 3, Dela. A 1951

- WEISSKOPF JOELSON, Edith
Dept. of Psychology (Ph.D.)
Purdue University A 1943
Lafayette, Ind. F 1951
- WELLS, Frederick Lyman (Ph.D.)
87 School Street H.M. 1950
Belmont 73, Mass.
- WELLS, Hal M.
66 Durland Ave.
Fluira, N.Y. A 1960
- WENGATE, Pauline (Ph.D.)
2321 Crescent Avenue
Charlotte 7, N.C. A 1950
- WERNER, Mr. Henry Clay
204 Eighth Ave.
New York 11, N.Y. A 1958
- WERTHEIMER, Rita (Ph.D.)
5500 Fieldston Rd.
Riverdale 71, N.Y. A 1955
- WEXLER, Rochelle M.
Apt. 6 D
130 Pelham Road
New Rochelle, N.Y. A 1949
- WHITE, Mrs. Helen Cecelia
1025 Worsham Drive
Whittier, Calif. A 1950
- WHITMAN, Mrs. Dorothy
3905 Winding Way
Cincinnati 29, Ohio
St. Aff. 1956
- WHITMAN, Roy M. (M.D.)
3905 Winding Way
Cincinnati 29, Ohio A 1954
- WHITSELL, Leon J. (M.D.)
52 Shore View Avenue A 1942
San Francisco 21, Calif.
- WICKERSHAM, Francis M.
U.S.P.H.S. Hosp. (Ph.D.)
Fort Worth 1, Texas A 1952
- WIGDOR, Blossom T. (Ph.D.)
4552 Lacombe Avenue
Montreal 26, Quebec, A 1949
Canada F 1956
- WILCOTT, Johanna Becker
(Ph.D.)
3900 Glenwood Rd.
Cleveland 21, Ohio
- WILDE, Dr. Guido
Apartado Aereo 11228
Bogota DE 2
Columbia, S.A. A 1955
- WILFENSKY, Harold (Ph.D.)
18 Essex Place
Hartsdale, New York A 1960
- WILKINS, Mrs. Verna M.
Mother Goose Nursery School
9500 Warren Street
Silver Springs, Md. A 1950
- WILLIAMS, Gertha (Ph.D.)
17211 Buckingham Dr. A 1944
Birmingham, Mich. F 1949
- WILLIAMS, Helen E. (Ed.D.)
253 West 72nd St., Apt. 1808
New York 23, N.Y. A 1950
- WILLIAMS, Mrs. Jessie M.
Griffins
Abinger Hammer
Nr. Dorking, England A 1950
- WILLIAMSON, Miss Margaret O.
350 Richmond Terrace
Staten Island 1, N.Y. A 1945
- WILSON, Helen Elizabeth
Eastern Montana
College of Education
1500 North 30th St.
Billings, Mont. St. Aff. 1958
- WILSON, Mary T.
R.D. 1, Box 94
South Salem, N.Y. A 1944
- WINER, Harold R. (Ph.D.)
8215 Westchester Drive
Dallas 25, Texas A 1956
- WOLF, S. Jean (Ph.D.)
220 Fifth Avenue
New York 1, N.Y. A 1944
- *WOLFSON, Mrs. Ruth
124 W. 79th Street
New York 24, N.Y. F 1940
- WOLPE, Zelda S. (Ph.D.)
435 North Bedford Drive
Beverly Hills, Calif. A 1950
- WOLTMANN, Adolf G.
1364 Lexington Ave.
New York 28, N.Y. A 1949
- WOOD, Austin B. (Ph.D.)
810 E. 19th Street
Brooklyn 30, N.Y. A 1943
- WOOLF, Henrietta K.
3345 Dent Place, N.W.
Washington 7, D.C. A 1950
- WRAY, Beulah K. (Ed.D.)
40 E. 10th Street
New York 3, N.Y. A 1949
- WRIGHT, Erik M. (M.D.)
Department of Psychology
University of Kansas
Lawrence, Kans. A 1943
- WRIGHT, Morgan
Winnipeg Gen. Hosp. A 1955
Winnipeg, Manitoba, Can.
- WRIGHT, Rogers H.
420 E. Carson St.
Long Beach 7, Calif. A 1961
- WUNDERLIN, Robert J.
Special Education Division
County Building, 4th Floor
Kalamazoo, Mich. St. Aff. 1961
- WYATT, Frederick (Ph.D.)
1027 E. Huron Street A 1948
Ann Arbor, Mich. F 1949
- WYLIE, Alexander
Box 902
Norfolk State Hosp. St. Aff. 1956
Norfolk, Neb. A 1957
- YADOFF, Bernard
5525 Bartlett St.
Pittsburgh 17, Pa. A 1958
- YANG, Andrew T. (Ph.D.)
106 Eileen Drive
Pontiac, Michigan A 1957
- YEAGER, Marion B. (Ph.D.)
3777 Norfolk 2A
Houston 27, Texas A 1960
- ZAMORSKI, Emil J.
719 Sheryl Drive
Pontiac, Mich. St. Aff. 1961
- ZEEV, Bracha
553 McDonald Ave.
Brooklyn 18, N.Y. A 1959
- ZFICHNER, Abraham M. (Ph.D.)
57 Northside Road
New Haven, Conn. F 1955
- ZIERER, Ernest (Ph.D.)
42-05 Lavton Street
Elmhurst, L. I., N. Y. A 1952
- ZIMET, Carl N. (Ph.D.)
333 Cedar Street
New Haven, Conn. A 1959
- ZIMMERMAN, Irla Lee (Ph.D.)
607 Bank of America Bldg.
Whittier, Calif. A 1949
- ZUCKER, Luise J. (Ph.D.)
276 Riverside Drive A 1945
New York 25, N.Y. F 1950

GEOGRAPHICAL DIRECTORY OF MEMBERSHIP OF THE SOCIETY FOR PROJECTIVE TECHNIQUES UNITED STATES AND TERRITORIES

ALABAMA

Tuscaloosa

Paul, Waters C.
Peyman, Douglas A. R.

ARIZONA

Nutt, Frances

Phoenix

Brewer, Paul W.
Canter, Aaron H.
Iverson, Norman
Katz, Harriet
Sargent, S. Stansfeld

ARKANSAS

Himmelstein, Philip

CALIFORNIA

Bowdlear, Charles
Bradway, Katherine

Cassel, Russell
Cole, Joseph
Grile, Marv
Due, Floyd O.
Everett, Evalyn G.
Gering, Evelyn
Klopper, Bruno
Krasner, Leonard
Lebowitz, Anne
Lewis, Robert T.
Payne, David H.
Sanders, David H.
Stanford, Margaret
Steen, Thomas
Sterne, Spencer
Tenney, Edward V.
Ullman, Leonard

Altadena

Brozovich, Stanley M.
Weir, John R.

Berkeley

Bell, John E.
DeVos, George
Inman, John M.
Miller, Christine
Reichard, Suzanne
Sundberg, Norman

Beverly Hills

Brandt, Rudolph J.
Dunn, Michael B.
Goldstein, Fred
Hansen, Irvin
Joel, Walther
Johnson, Theresa
Korot, Leonard
Lawrence, Ernest
Olinger, Leonard
Ruja, David H.
Simkin, James S.
Spiegelman, J. Marvin
Wolpe, Zelda S.

El Monte

Fracon, Helen
Malm, Mildred

Long Beach

Smith, Frances
Wright, Rogers H

Los Angeles

Armon, Mary Virginia
Beale, Elizabeth
Benson, Gerald
Bosker, Helga
Buckner, Charlotte
Carter, Louise
Crumpton, Evelyn
Dryden, Harold
Edison, Bernice T
Fanner, Betty G.
Fuchs, Neil
Hans, Kay B.
Farberow, Norman L.
Feifel, Herman
Fichman, Lionel
Foster, Berttram R.
Foster, Lucille K.
Frost, Marianne
Grayson, Harry
Hays, Berta
Hosmer, Evelyn
Howard, J. Stephen
Kay, Mrs. Victor
Knapp, Pearl G.
Lakin, Harriet A.
Lucas, Winifred
Makumzyk, Walter
Manson, Morse P.
Marsh, James
McDonald, Franklin R.
Meyer, Mortimer M.
Miller, Cecil R.
Mundaca, Harvey
Pellat, Marjorie
Raptin, Maurice
Rusch, Frank
Rose, Nicholas
Russell, Howard
Salzman, Anne
Shneidman, Edwin S.
Sommers, Vita Stein
Stendel, Kathleen
Troup, Evelyn

Palos Verdes Estates

DeVault, Helen C.
Holt, James
Struthers, Alice Ball

Pasadena

Diamond, Florence
Korda, Geraldine J.
Marsh, Donald
Price, Marian
Reitzell, Jeanne M.

San Bernardino

Peak, Horace M.
Plittman, Jack C.

San Diego

Brownell, Rosa Parsons
Davenport, Beverly
Heisler, Verda
Lockwood, Wallace V.
Stone, Irving R.
Treat, Wolcott C.

San Francisco

Berliner, Hildegard
Frankel, Esther
Harris, Robert E.
Kalis, Betty Lee
Meyer, George
Northcott, Hollie
Sanderson, Herbert
Whitsell, Leon

San Jose

Dunlap, Dorothy
Levy, Ruth J.
Mehr, Helen Margulies
Muench, George

San Marino

Snowden, Robert F.
Watkins, Roberta

San Rafael

Bergstrom Borland, Ingrid
Weinstein, Marvin

Santa Monica

Baker, Gertrude
Richards, T. W.
Sheehan, Joseph

Sherman Oaks

Pottharst, Karl
Reisel, Jerome

Whittier

White, Helen Cecelia
Zimmerman, Irla Lee

Woodland Hills

Cartwright, Robert W.
Reitz, Edna M.

COLORADO

Denver

Bernstein, Lewis
Billings, Edward G.
Fehrenbach, Alice R.
Kahn, Marvin
Malier, Alvin
Rogers, Lawrence S.
Ryder, Charles A.

CONNECTICUT

Barbara, Peter Paul
Fein, Leah
Gibson, Robert L.
Hemmendinger, Larry
Holmes, Frances B.
Holzberg, Jules D.
Kissinger, R. David
Molish, Herman
Rickers-Ovsiankina, Maria
Zeichner, Abraham

Bridgeport

Krall, Vita
Rosner, Stanley

Hartford

Nicholas, Alma L.
Olin, Tom Davis

New Haven

Ames, Louise Bates
Hellersberg, Elizabeth
Klatskin, Ethelyn H.
Zimet, Carl N.

DELAWARE

Holanchock, George

Wilmington

Blessing, Harold D.
Capaldi, Betty F.
Grossman, Searles
Weintraub, Irwin G.
Weiss, Sheldon

FLORIDA

Brodie, Dorothy B.
Gessner, Alan
Kelsey, Howard
Radtko, William
Seiler, Geraldine
Simkins, Lawrence
Stotz, Marion

Coral Gables

Chavkin, Albert
Lesser, Erwin
Spirer, Jess

Gainesville

Barry, John
Guertin, Wilson H.
Schumaker, Audrey S.
Schumaker, Henry C.

Miami

Allen, Robert M.

Eber, Martin
Karon, Samuel
Kushner, Malcolm
Marquit, Sybil

Tampa

Blau, Theo.
Schaffer, Robert E.

GEORGIA

Hughes, Robert M.
Kaplan, Norman
Sisson, Boyd

HAWAII

Honolulu

Dean, Sidney I.
Halperin, Sidney I.

ILLINOIS

Barrell, Robt. P.
Klass, Walter K.
Leiden, Irving
Neuman, Gerald G.
Shulman, Harold
Tabin, J. K.
Thornton, Thomas F.
Verrill, Bernard

Chicago

Altman, Charlotte H.
Beck, Samuel J.
Benstein, Hilde R.
Farley, Jane
Fromm, Erika O.
Handel, Gerald
Henry, William E.
Levy, Sidney J.
Lundin, William
McFarland, Robert L.
Moore, Harriet Bruce
Rainwater, Lee
Rosenthal, Vin
Schaw, Louis C.
Somerville, Addison
Surch, Jack E.
Swift, Joan Woodcock
Vassiliou, Mrs. Vasso

Evanston

Hilkevitch, Rhea R.
Stavrianos, Bertha
Vayhinger, John M.

INDIANA

Exner, John
Schmalzried, Newell
Indianapolis
Fortier, Robert H.
Graves, Winifred S.
Lewinsohn, Peter M.
Stoops, Wanda Rah

Lafayette

Baker, Lawrence M.
Weisskopf-Joelson, Edith

IOWA

Dingman, Paul R.
Pena, Cesareo D.

KANSAS

Mueller, Adolph R.

Lawrence

Kaplin, Bert
Wright, Erik M.

Topeka

Burdick, Jeanne
Mayman, Martin
Murphy, Lois Barclay

KENTUCKY

Johnson, Elizabeth
Louisville
Klinger, Marjorie
Vandenberg, Steven G.

LOUISIANA

New Orleans

Crovetto, Lorraine
Font, Marion McKenzie
Fosberg, Irving A.
Odom, Charles L.

MAINE

Doak, Mrs. James
Tyrell, Mrs. Marcel

MARYLAND

Abrams, Julian
Finn, Michael
Groff, Marne L.
Isaacs, Mark
Kensig, Isabelle V.
Lyon, W. B.
Rafman, Irving
Wilkins, Verna M.

Baltimore

Ainsworth, Mary
Feldberg, Theodore M.

Bethesda

Hoch, Erasmus L.
Little, Kenneth B.
Mindlin, Dorothee F.
Saltzman, Sara

Catonville

Imre, Paul
Shah, Saleem A.

MASSACHUSETTS

Evans, John T.
Kates, Solis L.
Pearce, Robt. F.
Stanton, Harriet
Tallent, Norman
Waller, Patricia

Belmont

Frank, L. K.
Wells, Frederic

Boston

DeVault, Barbara
Gilbert, Raymond R.
Low, Howard
Prentice, Norman

Brookline

Friend, Jeannette G.
Korner, Anneliese F.

Cambridge

Murray, Henry
Roe, Anne

Stockbridge

Harris, Robert A.
Philip, Anthony
Shapiro, David

Worcester

Lebeaux, Thelma
Phillips, Leslie

MICHIGAN

Brownfain, John J.
Capell, Martin D.
Fuller, Gerald B.
Guy, William
Klein, Louis
Leonard, A. T.
MacBride, John L.
Walton, Norma R.
Warren, Lurene Z.
Wunderlin, Robert J.

Ann Arbor

Blum, Gerald
Hand, Mary Ella
Hutt, Max L.
Schneider, Stanley F.
Watt, Frederick

Birmingham

Briskin, Gerald
Williams, Gertha

Detroit

Barahal, George D.
Eglash, Evelyn
Feinberg, Henry
Longley, James L.
Mather, Elise D.
Mathews, W. Mason
Milstein, A. Freda

East Lansing

Mills, David H.
Rabin, Albert

Pontiac

Yang, Andrew T.
Zamorski, Emil J.

MINNESOTA

Rotman, Saul
Seltzer, Samuel

Minneapolis

Rosen, Ephraim
Sulzer, Edward S.

St. Paul

Dossen, Herbert Jr.
Hernes, Christina
Scher, Sam C.

MISSOURI

Boutree, David
Cord, George A.
Heller, Arnold H.
Rook, Leroy
Waters, Thomas J.

St. Louis

Barclay, Allen
Estrada, Carol
L'Abate, Luciano
Osorio, Abel Garcia

MONTANA

Davison, Arthur H.
Wilson, Helen

NEBRASKA

Weiss, Bertram

Lincoln

Donskate, John
Jones, Marshall
Levine, David
Ruiz, Rene Arthur
Taulbee, Earl S.

Norfolk

Hillson, Joseph
Pesetsky, Fred
Wlie, Alexander

Omaha

Garfield, S. L.
Haworth, Mary R.

NEVADA

Dowlen, Caroline
Magnetie, Jules

NEW HAMPSHIRE

King, Francis W.

Manchester

Harmes, John
St. Clair, Walter F.

NEW JERSEY

Bosner, Jane Potter
Fabrikant, Benjamin
Feldman, Irving
Frev, Harriet K.
Gasorek, Kathryn
Greenstein, Jules
Kaufman, Bette
Krass, Alvin
Lawrenson, Thomas
Liutkus, Stanley
Morrow, J. Lloyd
Parnickv, Joseph J.
Roa-Che, Miriam
Selig, Kalman

Fast Orange

Abramson, Leonard
Rockberger, Harry
Tomblen, Donald

Englewood

Brody, Claire
Cohn, Ruth C.
Junkens, Elizabeth
Lieben, Beatrice

Maplewood

Goodman, Morris
Kutach, Samuel B.
Pecharsky, Adeline R.

Newark

Cune, Vivian
Steiner, M. Elizabeth

Princeton

Phillips, Maurice
Tomkins, Silvan S.

Trenton

Clark, W. Donald
David, Henry P.
Lyles, William

Upper Montclair

Brower, Daniel
Brower, Judith

NEW MEXICO

Cradick, Ray
Gersten, Rev. Chas.
Shackette, Sarah F.

Gallop

Mariani, Eugene
Mariani, Rose

NEW YORK

Abel, Theodora M.
Amich, Abraham
Appell, Melville
Benstein, Mikfred
Braun, Roslin
Chamberlain, Allan
De Martino, Hugo
Fron, Leonard D.
Friedman, Howard
Gurlich, Bernice M.
Herman, Jack
Hirsch, L. C.
Hirsch, Janet E.
Jankov, Mary
Kaplan, Herbert
Koss, Walter
Klein, Beatrice
Kotzsch, Morris
Kotzsch, Leonard
Lamp, Henry
Lewinsky, Julius
Levine, Abraham
Levine, Harold
Levinson, Boris M.
MacCasland, Mrs. Barbara
Margolis, Muriel
Napoli, Peter J.
Puzin, Frank S.
Ruswick, Serena
Seltzer, Samuel M.
Sharpe, Susie
Sobel, Albert L.
Wells, Hall M.
Wertheimer, Rita
Williamson, Margaret
Wilson, Mary T.

Astoria

Holodnak, Helen
Stern, Kathryn

Bronx

D'Angelo, Rita
Goodman, Pava
Harris, William W.
Iakowitz, David
Shor, Joel
Springer, Florence E.

Brooklyn

Abrams, Elias N.
Anderson, Helen Joan
Beurick, Milton
Dick, Frances
Hinds, Edith A.
Jortner, Sidney
Kavkewitz, Henry
Klein, Milton
Levinger, Leah
Lolis, Kathleen
Parker, Roland
Sitzman, Daniel

ORLANDO

Moss, Gerald
Norman
More, Alfred
Truman, Maurice

Orlando
Zachary, Anna
Wasson, Carroll B
Wardlaw, Norman

Portland

Coccarillo, Jennie
Good, Charles
Kerchner, Robert
Heggen, Gordon A
Klopf, Walter G
Mortson, Bernard
Bred, Max
Scott, Edward M

PENNSYLVANIA

Albright, Ray H
Barnett, Louis
Barber, Herman
Bricklin, Patricia
Cape, Eugene
Claus, Helen
Cox, Rachel D
Cummings, C. Peter
Decker, Robt J
Dougherty, Margaret
Feltz, Marion R
Ferar, Edgar
Hallow, William C
Heath, Douglas H
Johnson, Thomas
Kessler, Mable G
Kudorf, Irwin W
Levi, Herbert
Morris, Charles M
Ory, David
Phillips, John G
Platt, Henry
Rappaport, Sidney
Rauh, Edwin S
Ruhl, Mrs R. Ernest
Schulman, Irving
Schwartz, Lita
Singer, Roland H
Smolinsky, Harold J
Strat, Bennett
Taylor, Verda
Weiss, Evelyn R.

Penn. Mass

Beatty, Eleanor
McBride, Katherine

Penn. Mass

Jeffries, Helen
Morgan, Clellan L.
Morgan, Olive J.

Penn. Mass

Framo, James L.
Rappaport, Sheldon
Sless, Bernard

Penn. Mass

Alexander, Herbert M.
Kehan, Morton

Philadelphia

Abrams, Jules C.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

FOREIGN COUNTRIES

BRAZIL

Lopes, Jose Leme

CANADA

Clerk, Gabrielle Brunet
Ferguson, Kinsley
Francoeur, Thomas A
Howard, J. W.
Innes-Smith, Dr. J.
Leves, Arthur
O'Reilly, P. Oliver

BRITISH COLUMBIA

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.
Baron, Louis K.

ARGENTINA

Schlesinger, Alicia de

AUSTRIA

Cook, Philip H.

AUSTRIA

Forrest, Carol W.
Schoenhorn, Elizabeth
Simon, Maria D.

Record, Fr. Maurice
Spindler, Joan Elizabeth
Spies, Alan
Wright, Morgan

Montreal

Ariza, Fern
Beauchemin, Jean M.
Calvert, Margaret
Coughlin, Anne
Dudek, Stephanie
Greenberg, Nathan

Halpern, Esther
 Hebert, Bernard
 Kral, V. Adalbert
 Lehmann, Heinz
 Malloy, Helga
 Manuge, Mrs. Robt.
 Morf, Gustav
 Poser, Ernest George
 Wigdor, Blossom

Toronto

Carson, Marjorie
 Levinson, Mrs. Toby

COLUMBIA**Bogota**

Gaviria, Alvaro
 Rodriguez-Valderrama, Jose
 Wilde, Guido

CUBA

Lasaga, Jose

ENGLAND

Williams, Jessie M.

London

Alcock, A. Theodora

Bene, Eva
 Phillipson, H.

FRANCE

Cottes, Mlle. S.
 Morize, Mrs. Andre
 Schachter, M.

GERMANY

Hiltmann, Hildegard
 Rainwater, Mrs. H. B. Moore
 Vogel, Horst

INDIA

DaCunha, Dr. M. C.

ISRAEL

Guiora, A. Zeev

Tel Aviv

Akavia, Uriel
 Elizur, Abraham
 Kadinsky, D.

ITALY

Ferracuti, Franco

JAPAN

Hoshino, Akira
 Kataguchi, Yasufumi
 Motoaki, Hiroshi
 Tada, Haruo
 Tanaka, Fujio

MEXICO

DeDiaz, Nelly
 DeGovia, Guillermo
 DeLenero, Estela
 Fernandez, Rosa
 Nunez, Rafael

PERU

Sanchez-Garcia, Jose

PUERTO RICO

Dieppa, Jorge J.
 Umpierre, Francisco

SWITZERLAND

Cryns, Gerd M.
 Klurfeld, Georges
 Rorschach, Olga

JOURNAL OF PROJECTIVE TECHNIQUES

VOL. 26

JUNE, 1962

No. 2

. . .

EDITOR

BRUNO KLOPPER

EXECUTIVE EDITOR

BERTRAM R. FORER

ASSOCIATE EXECUTIVE EDITORS

WALTHER JOEL
JEANNE M. REITZELL
EDWIN S. SHNEIDMAN

STAFF EDITORS

EVELYN CRUMPTON
FLORENCE DIAMOND
NORMAN L. FARBEROW
HERMAN FEIFEL
FRED J. GOLDSTEIN
JAMES HOLT

WALTER G. KLOPPER
LEONARD KOROT
LEONARD B. OLINGER
NICHOLAS ROSE
JOHN WEIR
IRLA LEE ZIMMERMAN



The Color-Shading Response and Suicide¹

STEPHEN A. APPELBAUM AND PHILIP S. HOLZMAN
The Menninger Foundation

INTRODUCTION

Hunch or impression, bolstered by persistent efforts to verify it, often leads to fruitful advances in understanding. In the present study, we attempt to determine the validity of a clinical hunch.

For many years, several of us working with psychological tests at The Menninger Foundation have had the impression that a particular combination of shading and color² determinants of responses on the Rorschach Test is associated with suicidal tendencies in patients.³ None of us was bold enough to base any pronouncements of suicidal tendencies on the appearance of this one sign because we were averse to a blind sign approach to tests. But, as we continued to amass testing experience, the color-shading determinant appeared with insistent regularity in cases of patients who had attempted suicide, sometimes in the absence of any other alerting sign in a battery of tests. Curiosity about this consistency led us to test this hunch by systematic search.

The responses under consideration combine color and shading as determinants whether form is a primary or a secondary determinant. In scoring the Rorschach Test, following Rapaport's notation (Rapaport, *et al*, 1946), the responses were denoted as FC(C), F/C(C), C(C)F, C(C)/F, or C(C). As used here, (C) refers to

shading within the colored area, and not to achromatic color, scored C', or the shading to delineate areas, scored (C) on achromatic cards. It is analogous to shading in achromatic areas, scored Ch by Rapaport.

Some examples of these responses are:

1. Card 8, lower center half: "Shape and color of an iris; the velvety insides because of the shading." Scored FC (C).
2. Card 8, upper center blue: "Bluish crocheted material, nubby knit or boucle." In inquiry, the texture was attributed to the shading. Scored C(C)F.
3. Card 10, center pink detail: "Some kind of tissue, skin; it's burned at the edges." Inquiry brought out both color and shading as determinants. Scored C(C).

Over several years of unsystematic observation, the impression grew that regardless of the degree of morbidity of the response's content the color-shading determinant was the crucial variable indicating suicidal tendencies.

METHOD

If the color-shading determinant and suicidal proclivity are associated, the color-shading combination occurs more often among people who either attempted or committed suicide subsequent to testing, or who unsuccessfully attempted suicide prior to testing, than among people who made no suicidal attempts. To test this relationship, Rorschachs were gathered from seven groups of subjects: two experimental and three control groups matched on age and IQ, and two control groups not matched on age and IQ.

¹ We gratefully acknowledge the helpful criticism of several colleagues, among whom are Robert R. Holt, Roy Schafer, and Seward Hiltner. An abbreviated version of this paper was read at the 1961 meetings of the American Psychological Association, New York, New York.

² This observation was originally called to our attention in 1946 by Dr. Michael Dunn at Winter V.A. Hospital, Topeka, Kansas.

Experimental Groups

Suicides. Thirty Ss, seventeen female, thirteen male, who subsequent to being tested, committed suicide. Twenty-nine were patients in various psychiatric hospitals.³ Their psychiatric diagnoses included four neurotic depressions, seven character disorders, one involutional depression, two adjustment reactions of adolescence, one obsessive-compulsive neurosis, and nine schizophrenic reactions. One was not a psychiatric patient, but a subject in another study. Five subjects were patients in another hospital and their diagnoses were not available. The mean age was 33; the mean IQ was 120.⁴

Suicide Attempts. Thirty-nine Ss, thirty-two female, seven male, who made at least one suicidal attempt prior to being tested. All were patients in psychiatric hospitals. Their diagnoses included twenty character disorders, five neurotic depressions, four psychotic depressions, and nine schizophrenic reactions; one subject's diagnosis was not available to us. Because physicians differ in what they are willing to call an attempt, a histrionic display, or a gesture, our criterion was simply that the patient had made a motoric act of a suicidal nature. Consequently, the suicidal attempts of the patients in this group varied considerably in their appearance of seriousness. It is a safe assumption, however, that all patients in this group knew that their actions may well have been lethal. We did not include in this group patients who thought about suicide, whether fleetingly or profoundly, if they did not act upon these thoughts. The mean age was 32; the mean IQ was 119.

³ We wish to thank Dr. Arthur Kobler of Seattle, Washington for allowing us to study the Rorschach protocols he obtained in a West Coast Hospital from patients who subsequently committed suicide and from patients who had made a suicidal attempt.

⁴ All IQ's were determined from the Wechsler-Bellevue Intelligence Scale.

Control Groups

Non-suicidal Psychiatric Patients. Ninety-six patients, forty female, fifty-six male, who never made suicidal attempts. These cases were randomly selected from the records of the C. F. Menninger Memorial Hospital. The mean age was 35; the mean IQ was 117. Their diagnoses included thirty-seven character disorders, fourteen neurotic depressions, five psychotic depressions, twenty schizophrenic reactions, eleven anxiety reactions, six obsessive-compulsive neuroses, two adjustment reactions of adolescence, and one chronic brain syndrome.

Patrol. Fifty-three members of the Kansas Highway Patrol, all male, the "normal Ss" used in the diagnostic testing studies by Rapaport, *et al* (1946). These Ss are of a lower socioeconomic status than the other Ss in our study. At the present time, none of the Patrol has committed or attempted suicide. The mean age was 35; the mean IQ was 116.

Thyroid Ss. Fifty-two Ss, all female, in a study investigating thyroid gland dysfunction in women between the ages of 19 and 44. Twenty-four of these fifty-two were chosen from a midwestern city of 100,000 population by a stratified sampling technique. All twenty-four were without demonstrable medical or surgical pathology. Thirteen others were volunteer Ss for the thyroid study. These thirteen were also without significant medical or surgical pathology. The remaining fifteen Ss were referred to the thyroid investigation by physicians who suspected thyroid dysfunction. All subjects were given a battery of psychological tests and interviewed by a senior psychiatrist at The Menninger Foundation. The mean age was 34; the mean IQ was 115.

Psychiatric Residents. Fifty physicians, two female and forty-eight male, who had been accepted in the Menninger School of Psychiatry and who were tested as part of the school's

selection procedure. The mean age of 29 was significantly younger than all except the woman college student group. The mean IQ was 129 which is significantly higher than all other groups.

College Students. Seventeen students, all female, between 19 and 21 years of age, youngest of all the groups. They were originally Ss in an investigation of cognitive functioning. Wechsler-Bellevue Intelligence scores were not available for these subjects, but we presumed that they were at least of bright normal intelligence.

The Rorschach Test responses of all subjects in all experimental and control groups were read and scored separately by each author. The authors then compared their judgments and reconciled their differences in the few cases of disagreement. One author was unaware of what groups the test protocols were drawn from while he was scoring. The color-shading determinant was scored when the subject specifically mentioned the shading as well as the color as a determinant, or when the scorers were satisfied from the content that shading had been used, but the subject had not been specific in verbalizing it or inquiry was not adequate to elicit an unequivocal statement by the subject ("soiled wallpaper" in the mottled green area of Card IX).

RESULTS

Multiple appearance of the determinant was no more often associated with suicide than single occurrence ($\text{Chi-square}=0.224$). Therefore, the single occurrence of the sign was given the same weight as multiple occurrence. Table I shows the incidence of the color-shading determinant in the suicide, attempted suicide and control groups. There is a greater incidence of the sign among the successful suicidal patients (9 hits for every miss), than among the suicide attempt groups (4.5 hits for every

miss). In both suicidal groups, moreover, the sign occurs with greater than chance frequency, while in the control groups it occurs with less than chance frequency. A Chi-Square Test indicates that the difference between the combined experimental groups and the combined control groups could occur by chance less than once in a thousand times ($\text{Chi-square}=111.76$). These results obtain in spite of the as yet unknown effect of the high IQ and lower age of the residents, and lower age of the college students.

While one might have anticipated that the experimental groups would show fewer color responses with form as the primary determinant (FC) than the control groups, all groups used C(C)F proportionately more often than FC(C) and C(C), with no significant difference between the combined experimental and combined control groups ($\text{Chi-square}=2.417, p<.30$).

The incidence of color-shading responses with "morbid" content (e.g., "decaying flesh") is greater in the combined experimental groups (seven out of 69 color-shading responses) than in the combined control groups, including non-suicidal patients (two out of 51 color-shading responses). The low frequency in all groups of such "morbid" responses, however, indicates that the content of the color-shading responses is not a

TABLE I—Incidence of the Color-Shading Determinant in Suicidal and Non-Suicidal Subjects

		Color-Shading	
		Present	Absent
Experimental Groups	Suicides	27	3
	Attempted Suicides	32	7
	Sum	59	10
	Non-Suicidal Patients	18	78
Control Groups	Thyroid Study Subjects	9	43
	Highway Patrol Subjects	5	48
	Psychiatric Residents	12	38
	College Students	7	13
	Sum	51	220

good distinguishing measure between the experimental and control groups.

For both experimental and control groups the color-shading responses appeared in descending order on VIII, IX, X, II and III, perhaps on the basis of the structure of the card.

Thus, the significant link is between suicide or attempted suicide and the simple presence of the color-shading determinant, irrespective of the number of instances in an individual protocol, how this determinant is combined with form, the morbidity of content, or on what card it appears.

Is the sign related to other determinants such as C and Ch or the number of responses (R)? Positive correlations were found in all groups between the color-shading determinant and the total number of responses per protocol ($r_{p, bis} = .31$, $p < .001$), presence of achromatic shading responses ($r_{tet} = .38$, $p < .001$), and number of color responses independent of those combined with shading ($r_{p, bis} = .58$, $p < .001$). Would these other scores, then, also successfully predict suicide? Forty-nine out of ninety-six nonsuicidal hospitalized patients gave at least one Ch response, and eighty-four of ninety-six gave at least one C response. The high rate of false positives invalidates their use as a sign of suicidal actions.

The effectiveness of the color-shading determinant as a predictor of suicide depends upon the sign's increased predictive power beyond what one would expect from a knowledge only of the base rate of suicide in the population for which one is predicting. To estimate the base rate of suicide in our population we selected a random sample of 252 other patients from admissions to the C. F. Menninger Memorial Hospital during 1958, 1959 and 1960. From the records of these patients we determined that twenty-eight per cent had at some time attempted or committed suicide. We took this figure of

twenty-eight per cent as the base rate of suicidal actions in our hospital patients. Table I indicates that the test sign correctly identifies eighty-six per cent of suicidal patients, a significant increase in predictive power beyond the twenty-eight per cent base rate. Use of the sign incorrectly labels nineteen per cent of the control population as potential suicides. For our hospital population the test sign is clearly an alerting signal of suicidal behavior. Absence of the sign in our hospitalized patients, however, is no valid indication that there is no danger of suicidal behavior. The fourteen per cent false negatives (10 out of 69) are evidence that the sign is not an invariable predictor, and that its presence in a record is a surer alert to suicidal action than its absence is a guarantee of no such action. The incidence of false positives in the control groups (nineteen per cent) however, exceeds our estimate of the base rate of attempts and suicides in the general population (approximately one percent)⁵, indicating that the sign is not an efficient predictor of suicide in the general population.

DISCUSSION

The color-shading determinant occurs with significantly greater frequency in groups of suicides and attempts than in five control groups. For our population, at least, predictions of suicide of hospitalized psychiatric patients made from the use of the color-shading determinant are considerably more effective than predictions made from the base rate of suicide in that population. But the determinant is also used by some patients (nineteen per cent) who do not attempt suicide, although it is likely that a few of these nineteen per cent false positives would make a suicidal attempt were they not hos-

⁵ There are no reliable studies of the rate of suicidal attempts. It seems to us unlikely that the rate of attempts in the general population would approach nineteen percent.

pitalized. The clinical records of several of these patients show reports of serious ruminations about suicide, but so do the records of those who do not give the determinant. It would be interesting to be able to see which patients subsequently make suicidal attempts: whether those patients who translate their suicidal rumination into suicidal action come from a population of "susceptibles" that may be identified by the color-shading determinants.

Among the control groups, however, the number of false positives far exceeds the death rate from suicide in the general population (approximately one per cent) and our estimate of the number of suicidal attempts in the general population (less than one per cent). Using this determinant as a predictor of suicide is defensible, therefore, only with hospitalized psychiatric patients. These data support the proposition that the process underlying the use of the color-shading determinant leads to suicide only when the person's total functioning becomes so impaired that he requires hospitalization.

What might be the significance of color-shading responses for adequately functioning people and for those whose personality functioning is impaired? That the determinant is more common among the psychiatric residents and college students than among our other control groups suggests that it may reflect a potential asset in well-functioning people of high intelligence, a speculation consistent with the fact that suicide is positively related to occupational and socio-economic levels.

The color-shading determinant seems to reflect an unusual differentiation in responsiveness to color in the Rorschach Plates. Rather than seeing the colored area as a unity—"blood because it is red"—the subject is sensitive to additional properties which make the response more individual—"dried blood because it is

red and the shading indicates it is encrusted." Rather than "blue pillows" he sees "blue satin pillows," the shading indicating the sheen of satin. Thus, there is a searching, going-beyond-the-givens quality; a spade is more than just a spade. But this approach is also used in giving shading responses to achromatic areas, and our data have shown that Ch is not a useful predictor of suicide. Two possibilities present themselves: (1) Since more people use achromatic-shading than color-shading as response determinants, it may be that insofar as the Rorschach Plates are constructed, color-shading is more difficult to use in a Rorschach test response. All people who can make use of color-shading can use achromatic-shading, but not all people who can use achromatic-shading may be able to use color-shading. Thus, people who use color-shading may possess the searching, penetrating characteristics hypothetically related to this response to a greater extent than those who do not, or cannot, use this color-shading determinant. (2) The way one responds to color on the Rorschach Test is assumed to reflect the person's idiosyncratic affect organization. One could, therefore, infer that the searching, highly differentiated quality of the color-shading response reflects a sensitivity to nuances of feeling. Rapaport, *et al*, stated that an accumulation of these color-shading responses "... suggests a person given to much mingling of affect and anxiety, as is seen in complex emotions like nostalgia (sweet-sorrow)." (Rapaport, *et al*, 1946)

Both of these explanations suggest that persons who give even one color-shading response have a keener eye for what is there. They are more aware of nuances, not only of the colors they see but perhaps of objects in general, and can see novelty where others see only the familiar. The "work" done in using the color-shading determinant consists in an articulating, discovering and penetrating

activity. In adequately functioning people of superior intelligence, such sensitivity and capacity for differentiated responsiveness may be available for adaptive uses, for example, in discriminations between ideas, deep inquiry into the uniqueness of objects, empathy for the nuances of another's experience, and the appreciation of art.

However, patients excessively concerned with their difficulties and caught up in their anxieties about living may turn the capacity to penetrate beyond the obvious toward an inquiry into their own existence, and the purpose of their own being. They may then ask what they want from life, whether and how they can get it, and whether the struggle is worth it. They face the problem of Sisyphus, who could endure the physical struggle of continuing to push the rock up the hill, but who was tortured by the recognition that it was all without purpose, for the rock immediately tumbled down again. The price of such near-sighted clarity is probably high for it seems to lead to a loss of perspective, a "centration," in Piaget's terms, upon his feelings in the immediate present. (Piaget, in press) Note Freud's insight into the heightened awareness of people in misery as he describes the puzzling astuteness of melancholiac patients: "... he has a keener eye for the truth . . . When in his heightened self-criticism he describes himself as petty, egoistic, dishonest, lacking in independence, one whose sole aim has been to hide the weakness of his own nature, it may be, so far as we know, that he has come pretty near to understanding himself; we can only wonder why a man has to be ill before he can be accessible to a truth of this kind." (Freud, 1917)

What Freud knew, but merely suggested in his faintly ironic, rhetorical question, is that this terrible clarity of detail issues from a distortion of one's perspective of himself in rela-

tion to his *entire* life situation. A too close look at a tree obscures the recognition that it is but one tree in a forest. Not only does the person see only one aspect of himself, but there is the implication that time too is restricted to the all-encompassing present. The past, when things were better, seems blotted out as does the future when things could be better. Gone with the sense of time are its palliative and protective functions. With the loss of perspective of self and history, the momentary situation of travail becomes everything. Various alternatives and possibilities of achieving change for the better seem unavailable. Data on suicide further support our belief that the experience hypothesized here leads to a short-sighted perspective of one's historical continuity and to a feeling that one has run out of alternatives: suicide rates decrease significantly during times of national crisis, when there are socially imposed purposes and goals, and the rates increase during times of surplus leisure and personal idleness. These patients would be better served if they were buffered by the refuge of greater generality, less involvement, and an increased ability to "let it go at that."

We do not suggest that there is only one kind of suicide, either from the standpoint of dynamic motivation or descriptively. Shneidman and Farberow (1957) have detailed varieties of suicides, and a study by Dublin and Bunzel (1933) shows the multiple roots of self-destruction. Presumably, suicides can occur as "pure" ritual acts or as a sensible alternative to intractable pain. Or, perhaps the experience of exquisite sensitivity reflected in the color-shading response may be present in admixture even in these situations, for there remains the question of why some people do not commit suicide under apparently the same conditions obtaining when others do. The state of mind we have suggested here to explain our results is complementary to motivational

theories of suicide (viz., Menninger, 1938). This state of mind provides what might be a structural condition, which, regardless of the precipitating motivation or circumstances bringing about the distraught condition, leads the person to view existence in such a way that suicide appears his only available course of action.

We offer these speculations as plausible maps to direct our search for the processes underlying the color-shading determinant and its relationship to suicide. For further information and greater degree of certainty, it is necessary first to seek replication of these findings from other psychiatric hospitals and then to find personality attributes that are associated with the appearance of this determinant. In the meantime, in populations of patients such as the one presented here, it seems that the presence of this sign may take its place in the total clinical examination as a significant bit of evidence pointing toward suicidal tendencies.

SUMMARY

The use of shading in colored areas of the Rorschach Test Plates by hospitalized psychiatric patients as an in-

dicator of suicidal tendencies significantly improved predictions of suicide or attempted suicide over those made from the base rate in that population. The color-shading determinant occurred significantly more often among the suicide and attempt groups than among five control groups. We offered speculations about psychological processes that may be involved in the color-shading determinant, and about some consequences of the assumed processes in suicidal and non-suicidal populations.

REFERENCES

- Dublin, Louis I. & Bunzel, Bessie. *To be or not to be*. New York: Harrison Smith and Robert Haas, 1933.
- Freud, Sigmund. Mourning and melancholia (1917) in *Standard edition of the complete psychological works of Sigmund Freud*. Vol. 14. London: Hogarth Press, 1957. Pp. 237-260.
- Menninger, Karl. *Man against himself*. New York: Harcourt, Brace and Company, 1938.
- Piaget, Jean. Will and action, *Bull. Menninger Clin.* 26: 138-145, 1962.
- Rapaport, David, et. al. *Diagnostic psychological testing*. Vol. 2. Chicago: Yearbook Publisher, 1946.
- Shneidman, Edwin S. & Farberow, Norman L. *Clues to suicide*. New York: Blakiston, 1957.

Received October 25, 1961

Psychometric Aspects of the Rorschach Technique

WILLIAM E. BLOCK
Yeshiva University

INTRODUCTION

A generation has elapsed since Hermann Rorschach first published his psychodiagnostic technique. Subsequently, the Rorschach technique has become the pre-eminent weapon in the clinician's armamentarium and has given rise to at least 2000 major and minor research studies. A body of beliefs or "clinical lore" has emerged out of the individual clinician's experience with the technique. Facility with the Rorschach technique has become a virtual hallmark of the clinical psychologist.

Yet recent laborers in this vineyard have harvested a bitter crop or two regarding projective techniques in general and the Rorschach in specific. Murstein (1960) among others has commented upon the refractoriness of the Rorschach technique to factorial experimentation. Masling (1960) has indicated that the Galtonian rationale regarding the objectivity of tests, the x-ray concept, no longer seems valid for projective techniques, responses to which are a function of mode of administration, testing conditions, and interpersonal factors operant during testing. It would seem that the area of projective techniques is one of the last bastions of classical scientific thinking in psychology to yield to the principles of indeterminacy and relativity characteristic of measurement in other fields.

Prior research with the Rorschach has largely been focused on *what* is measured, the meaning of various categories of response, with scant attention directed to the more fundamental question of *how* the Rorschach provides the raw data upon which subsequent interpretation depends. It would seem before all else that the

Rorschach is a psychometric experiment, that is, a more or less standardized stimulus situation eliciting certain responses from the subject. However, the psychometric rationale of the test has not been fully explicated since Hermann Rorschach was a gifted innovator rather than a skilled psychometrist. As with any measuring instrument, both "true" and "error" variance exist, except that in the Rorschach the proportions remain largely unknown. Herein lies the dilemma of the clinician. He can never be certain whether a Rorschach protocol represents a true apperception or revelation of the patient, or an artifact brought on by errors inherent in the basic psychometric procedure employed. Often, the errors of measurement are of the same magnitude as the differences sought by the clinician.

It is the author's hope, therefore, that this preliminary inquiry into the psychometric rationale of the Rorschach may serve to sharpen the clinician's awareness and suggest ways of transcending the present dilemma of the Rorschach.

The clinician may confront this dilemma in two ways, either by "bracketing" out of awareness paradoxical elements and pursuing his lonely idiosyncratic course on an intuitive basis or by admitting these questions into consciousness and by so doing transcend the limitations of his present view to new, stranger horizons.

Actually, there is little choice. For the relationship between psychometric rules and projective techniques is intrinsic like that between the rules of grammar and language, which operate, whether we are aware or not.

CLASSICAL PSYCHOMETRIC VIEW OF THE RORSCHACH

The classical psychophysics of Fechner, Müller and Wünderl concerning itself with quantifying the relation between the physical event and corresponding psychological events, or the relationship between stimuli and response. Implicit in this quest was the assumed existence of two quantitative variables, a physical (stimulus) continuum paralleled by a psychological (response) continuum. Each physical continuum consisted of measurable physical units representing a single change in some physical property corresponding to which there are certain aspects of sensory experience. The response continuum is shorter than the physical continuum since some stimuli are too weak to arouse any response and others too great for the sensory receptors. (Guilford, 1953).

Viewing the Rorschach in the light of classical psychophysics, it may be regarded as a psychophysical continuum of 10 visual stimuli paralleled by a response continuum. The upper and lower threshold features might be illustrated by the virtual impossibility of giving a whole response on card X because of its dispersion of detail; the fact that popular responses can be readily elicited; or that the meaning of card VI is unclear semantically to most subjects (Rabin 1959).

The Rorschach technique is an admixture of several psycho-physical methods, corresponding to its main phases or *free response*, *inquiry*, and a quondam third phase, *testing limits*. Its scoring and interpretation involve various levels of measurement.

The *free response* phase, in which the subject reports what he sees in the blot, appears to involve the use of constant methods, specifically that variant known as the method of single stimuli with variable standard. Thus, each card in the stimulus series of 10 cards, is presented sequentially in a prearranged manner such that each

card in turn becomes the standard against which others are subsequently viewed. This procedure is exemplified by S's comments about similarities or dissimilarities between blots, other adjectival qualifiers comparing blots, and S's tendency to respond to the cards holistically rather than discretely.

The inquiry phase is difficult to analyze, for more than one psycho-physical method seems to be involved, so that no one feature predominates. There are involved the use of successive categories, e.g. in location, determinants, the need for S to discriminate between pairs of stimuli (W,D,F, M,C,c) as in the method of paired comparisons, and the hierarchical assignment of determinants, one per category, as in the method of rank order. It would seem that the method of rank order best describes the process of the inquiry, for there is only one rather than successive categories for location and determinants, and all stimulus aspects of the blot are present for simultaneous observation unlike paired comparisons.

PSYCHOMETRIC METHOD AND RORSCHACH VARIANCE

The apparent psychometric strengths of the Rorschach have been frequently detailed in the literature, namely, the simplicity of response required of S, the sequential nature of responses that permits interaction or cumulative effects and the attempt to utilize a quasi-objective scoring system. This has not been the case with apparent psychometric weaknesses intrinsic to the Rorschach by virtue of the basic psychometric methods involved and the peculiarities of human judgment that influence the operation of certain methods of measurement. Such a detailed analysis will now be undertaken.

1. Constant methods allow for biased judgments to operate. Thus if responses to the Rorschach may be regarded as preferential, S may bias his responses in certain non-commit-

tal directions, e.g.—neutral, popular, or stereotyped responses as opposed to giving an emotionally disturbing response. While such responses may in themselves be revealing about S, the effect is to shrink the response continuum. What S's potential full range of response might be then becomes a matter of inference. Testing-limits was introduced to offset this effect. Studies cited by Masling (1960) show that Ss can constrict protocols under threatening motivational sets.

2. The constant method raises the question of symmetry. That is, the Rorschach series is essentially asymmetrical in that responses tend to drift towards a geometric mean of the stimulus values of the cards so that unknown constant error may exist in the time or space orders of successive cards. This "time-order" error may be of the type wherein the second of two successive cards or two successive responses gives a stronger impression compared with the first or sinking impression; or where the trace of the first stimulus assimilates to the level of excitation existing between the two stimuli compared. The "constant error" paradoxically *varies* with the general level of stimuli represented by each card as well as the series in toto, prior sets of S, background stimuli, suggestion and environmental conditions. Most of these conditions have been shown experimentally to influence Rorschach responses in one way or another (Masling 1960). Constant error effects cannot be neglected because of the reliance on sequence in presentation of the cards and in various aspects of interpretation. Since this type of error is not known in advance if at all, sequential analysis in the Rorschach may not be wholly defensible.

The method of rank-order in the inquiry tends to enforce upon S a level of internal consistency which may be more apparent than real and with resulting loss of degrees of freedom. The method is illustrated in whether F takes precedence over C, M

over C, and so on.

But the real difficulty inherent in the inquiry—if the method of rank order describes it properly—is that the ranking is on the basis of an ordinal scale. The logical basis of the ordinal scale is that the things ranked differ in some quality. Sometimes the basis of classification is a composite of two variables,—M is both form and movement. But in ranking complex responses on the Rorschach, S perforce superimposes a multidimensional quantity onto a linear scale—intuitively weighting component variables—so that E may never know what these weights are or whether S applied them uniformly.

Another problem in connection with ranking in the Inquiry is that the number of stimuli may be too large in terms of S's interest or capacity to discriminate. Blots that are more intriguing for S will tend to elicit greater discrimination.

Perhaps the most fundamental weakness associated with rank-ordering in the inquiry is the fact that S is simultaneously exposed to all stimulus properties of the blot in making the original response, adapting as it were to a composite standard—the pooled effect of all stimuli—but that in the Inquiry this complex behavior is treated unwarrantedly as a simple additive procedure.

Testing limits, which requires S to give certain responses—usually populars—not given in the free response phase, is a complicated procedure. On the one hand, it requires S to match a visual stimulus to a verbal standard, resembling the method of average error. On the other hand, the stimulus concept is made to approach the stimuli by successive addition of information by the examiner as required by S, resembling the method of limits.

JUDGMENTAL FACTORS IN RORSCHACH RESPONSES

So far, the author has attempted to describe the phases of Rorschach

testing via classical psychophysical models. The basic assumption of classical psychophysics was that of a univocal relationship of stimulus to response—a linear regression (with perfect correlation) relating S's judgment of what is experienced to the response *per se*. But there is no way for the experimenter to obtain direct evidence of the exact quantity of the response any more than the subject can produce these for inspection. The evidence that the subject provides the experimenter is a form of judgment whether verbal or symbolic. The evidence in psychometric studies is that the regression is not always linear nor the correlation perfect (Guilford, 1953).

What this portends for the Rorschach is this: a judgmental continuum intervenes between the visual percept and its verbal articulation; in other words, Rorschach responses are in the nature of intervening variables. One deals in the Rorschach with overt judgments from which one infers certain prior responses. The same point has been pithily expressed by Murstein (1960) who states, "The Rorschach is operationally hardly a perceptual test but rather an interpretation of an interpretation."

Thus in the free response phase, S makes a judgment with respect to the stimulus properties of the blot vis-à-vis a contemplated response; one response may be judged with respect to an antecedent response, or one card with respect to another; during inquiry, S falls back on introspective judgments vis-à-vis the free response; and in testing limits, S judges the similarity of a verbal concept and its physical stimuli counterpart as paired comparisons.

It would appear then that the general problems of psychophysical judgment are relevant to the Rorschach situation. This view is supported by the fact that psychometric judgments are influenced by focal, contextual and residual features of the experiment, the same type of influences

which have been skillfully adumbrated with respect to the Rorschach test situation. (Masling, 1960).

It remains to say that in addition to the constant errors one finds associated with those classical psychophysical methods, special influences on judgment such as sets and adaptation level probably operate with respect to the Rorschach too.

The method of limits embodied in testing limits yields several types of errors. There are constant errors of *habituation* that arise from the fact that S knows that successive cues from E are in the direction of increasing similarity between suggested verbal stimulus and the sought-after response. Moreover, since S knows too what is expected of him—namely, to see the suggested percept—this heightened attention and suggestion may lead S to report the percept before it is phenomenally experienced.

Furthermore, there are apt to be *stimulus errors*, in which certain cues emanating from the stimulating conditions may serve as irrelevant criteria for judging the percept. S is informed prior to limits testing that he will be told what other people see. Since S now knows that the suggested concept is supposed to be equal to the physical stimuli of the card being tested, he may take his two corresponding phenomenal experiences to be equal, which may not at all be the case given the imperfect correlation between the judgment and response continua.

One further major source of error in testing limits is the haphazard manner in which successive cues are given to S by E. This *haphazard presentation of stimuli* places the method on a hit-or-miss basis with uneconomical use of cues. It is of historical interest to note that Kraepelin purposely utilized this haphazard method in an effort to eliminate some sources of errors (Guilford, 1953). However, the substitution of unknown and thereby largely uncontrollable errors for known and perhaps controllable ones seems dubious procedure.

JUDGMENTAL SOURCES OF RORSCHACH ERROR

It was stated earlier that error variance in the Rorschach technique derived from the psychometric methods intrinsic to the test as well as judgmental components. The latter source of errors will now be considered.

An ostensible virtue of the Rorschach technique in comparison with other personality measures is its ambiguous, unstructured quality. S thereby is required to define his own task given minimal cues by E, and thus reveals idiosyncratic elements of his personality make-up. Schachtel (1945) among other Rorschach experts attaches interpretive significance to this subjective definition of the Rorschach situation. Paradoxically, this clinical virtue is a psychometric liability. If each S defines his own task or goal the stable conditions necessary for meaningful scoring and interpretation are largely indeterminate.

Moreover, the more enigmatic the task the more likely will S fall back on response biases and sets, so that responses other than those sought may predominate. That is, we are more likely to detect adaptive behavior than deep-seated personality traits. A number of studies (Masling 1960) have already appeared indicating how the mode of presentation of the test, and its definition by the subject can influence Rorschach profiles as well as specific determinants. However, as long as each S is free to define the task for himself, one can never judge the appropriateness or not of the manifest adaptive behavior.

It may help to identify some likely sets on the part of S in the Rorschach situation. Schachtel (1945) has identified an acquiescent set that operates in relation to S's perception of E as an authority figure. Cronbach (1946, 1950) discusses many kinds of sets occurring within various test situations, some of which have relevance

for Rorschach testing. The *semantic set* refers to the unique meanings test stimuli have for S; that such sets operate with respect to the Rorschach is verified in the work of Rabin (1959), who found wide individual differences in meaning of the blots on the semantic differential scale. Under the *gambling set*, S's responses will either be cautious or profligate depending upon his estimate of his chances of guessing correctly or not. If S is motivated by a *falsification set*, the desire to make a good showing, to cover up defects and deficiencies, his responses will be so tailored; Masling (1960) cites experimental evidence supporting the existence of a falsification set on the Rorschach.

According to Guilford (1953), response sets follow certain principles: 1) Sets are consistent within the individual and persistent in time; thus a set may inflate the reliability of final scores.

2) Sets confound scores. Whether the set adds to both true variance at the expense of common variance or just the former, the meaning of a score and its interpretation is placed in doubt.

3) Sets are a function of task difficulty. The more difficult is the perceived task confronting S, the more likely response habits and sets will be elicited. Since the Rorschach cards are of varying difficulty¹, with degree of difficulty varying with each S, variable response sets may be expected so that each protocol is bound to have its own degree of unreliability.

CRITIQUE OF RORSCHACH SCORING

Psychometrics is concerned with empirical data, one type of which may be test scores. Though Rorschach practitioners would prefer to stress interpretive and configurational aspects of the test, by virtue of the fact that these employ a formal scoring

¹Compare I and V with IX; the absence of general semantic meaning of VI (Rabin 1953)

procedure, the scoring rationale and procedure fall within the purview of psychometric theory. Moreover, even an "intuitive" approach to a test protocol assumes some scale or framework in which an interpretation can be sustained, and any interpretation purporting to mean something implies this scalar notion. The scoring problems of the Rorschach will be examined first as to basic measurement rationale, and secondly in terms of statistical practices.

The inherent measurement rationale of the Rorschach is primarily that of the nominal scale, and secondarily, the ordinal scale. The nominal scale serves merely to designate class or category. In this sense the arbitrary division of a perception into location, determinants, and content, and the numerical designation of the 10 card series are nominal in nature. Though a first step, in measurement is nominal classification, it limits the kinds of statistical manipulations possible. But even of greater concern is that the nominal scale requires identity or uniqueness of categories. Yet certain scoring categories are not independent, viz: human movement involves form, movement and human content. The criteria for classification, moreover, are to some degree ambiguous. A perception may be variously scored depending upon nuances in S's verbalization and the scoring system used (Beck, Klopfer). Also a perception involving a complex determinant or several determinants involves a hierarchical scoring order, wherein the determinant deemed most significant assumes priority over a lesser one, giving the appearance of greater frequency, viz: M vs. C, C vs. Fc. etc. As Murstein (1960) points out, discriminating interpretations require the number of responses to be independent of the frequency of determinants. In this rank ordering of determinants, as well as the crude sequence of cards finally established by Hermann Rorschach, we have elements of the ordinal scale.

But perhaps the basic weakness is the quasi-objective system of Rorschach notation whereby certain features are abstracted from each response. This shorthand notation is highly dependent upon S's powers of verbalization, language nuances, as well as the examiner's judgment of goodness of fit between stimuli and response. In addition studies cited by Masling (1960) show a host of iatrogenic influences on scoring subsumed under warm-cold examiner behavior, operant conditioning, the stimulus value of E as well as his personality characteristics. Thus superimposed on basic scalar weaknesses are those having to do with personal characteristics of S and E.

Certain assumptions are made regarding levels of measurement in actually handling of scores which appear unwarranted. Many of these have been touched upon here and there in the literature. Some further comments are also in order.

First there is the practice of summarizing additively the scores assigned each card to arrive at the psychogram. This involves the unproven assumption of an interval scale of measurement, such that each card is equidistant from its neighbors so that distances between cards I and II mean the same psychologically as between II and III, or IX and X. Unequal intercorrelations among cards make this assumption somewhat questionable.

The various complex formulae or ratios employed in the psychogram, such as M: Sum C, are interpreted as though they were in fact derived from a ratio scale. That is, the ratio scale has the property of absolute or genuine zero, where a zero score means none of the property being measured. This presents no problem so long as mere frequencies are dealt with, e.g. 10W, 5D, 3M, zero C. But the moment we interpret scores as indicating the position of the subject on a trait scale, i.e. high drive, lack of emotionality, lack of drive, etc.; we are treating a simple frequency as a ratio

score. Not having demonstrated that the Rorschach embodies ratio scale properties, we have no grounds for making absolute statements. Zero color responses may mean more than absence of affectivity; also 4M does not imply a proportional increase in creativity over 2M; nor that an M: Sum C ratio of 2:1 is half as significant psychologically as M: Sum C of 4:2.

The intrinsic weakness of the scoring system, however, lies in the variable nature of response productivity, R. Response productivity enters into many indices, percentages and ratios employed in the psychogram. But R has been found to be subject to widespread individual differences such that these may be carried over into specific categories as artifacts. The use of percentage values of R is misleading since some percentages will diminish with lengthier protocols because of lesser likelihood of some responses, viz. wholes and populars. Recent empirical findings (Anatasi, 1954) show that R varies also with age, intelligence, and education of the subject, as well as his verbal fluency, and from examiner to examiner.

Thus the keystone of the scoring edifice, R, rests on an unstable base of variables extraneous to those purportedly assessed by the Rorschach.

The alleged innate capacity of the subject to perceive form and movement responses has recently been shown to be dependent on linguistic aptitude, a product itself of educational influences (Anastasi, 1954).

A final comment on the statistical operations with Rorschach data involve assumptions of additivity. Additivity requires that $a + b = b + a$, that is, the order in which things are added makes no difference. This requirement is ostensibly met in computing Rorschach Psychogram statistics, for example:

$$\text{Sum C} = \text{FC} + \text{CF} + \text{C}$$

$$\text{FM} + \text{Fm} + \text{m} :$$

$$\text{FC}' + \text{C}'\text{F} + \text{Fc} + \text{cF}$$

It matters not at all how the terms are added. Yet distinct qualitative and quantitative differences are intended by the notation of FC vs. CF, Fm vs. mF, Fc vs. cF. So that in one order of operation the qualitative features are disregarded; but are elevated for other operations. Thus the basic postulate of additivity would appear to be used in contradictory senses, for under the latter conditions $a + b = b + a$.

Suffice it to say that the present scoring system, with its subjective elements, and its psychometric inconsistencies lead one to question the degree of stability of the resulting statistics.

COMMENTS ON RELIABILITY AND VALIDITY OF THE RORSCHACH

Implicit in the foregoing discussion are the complementary questions of reliability and validity of the Rorschach. The problems here are not unique to the Rorschach as a representative projective technique but are more complicated than in the case of the structured-objective type test. Some theoreticians, among them Rosenzweig (1951), while arguing for redefinition of these standard psychometric concepts as applied to projective techniques, do not eschew their relevance. The author will make no effort to deal with or summarize the many well-known views on this problem but rather to point up some basic questions.

Reliability in the Rorschach technique is a complex resultant of such elements as the mode of administration, transient needs of the subject, inherent psychometric features of the technique, the scorer-score interaction with its judgmental aspects, and the rules of interpretation applied by the scorer. To put it more succinctly, reliability is a resultant of the triadic interaction of subject-test-examiner. Unfortunately, at this stage of our knowledge, the precise nature of the interaction is not yet known sufficiently to assess the error contribu-

tions from each source. Though there are some steps that can be taken to minimize errors to be discussed shortly.

Paradoxically, factors adversely influencing reliability may contribute to the validity of the Rorschach. For example, response sets tend to lower reliability yet eliciting such sets may be a valid aspect of the Rorschach. Just what is uniquely valid in the Rorschach remains a moot question. Cronbach's review (1949) of Rorschach studies indicated about half of the experimentally tested hypotheses are consistent with theory, that low positive (.35) relationships were obtained between specific Rorschach indicators and postulated traits; and that some aspects of theory are patently false. The "broad band width" and "low fidelity" of the Rorschach—its global aspects as opposed to its pin-point accuracy undoubtedly account for the lack of clear-cut validation. Although other sources of error are the criterion groups employed, and whether or not these possess such desiderata as relevance, freedom from bias and reliability.

The Rorschach seems to have an inherent refractoriness toward standardization and norms, as Hermann Rorschach was among the first to point out. Given the plethora of non-definitive validation studies, the chief type of evidence marshalled in behalf of the validity of the Rorschach is the individual practitioner's experience with it in a clinical situation—the so-called "face validity."

Face validity, however, is *procedural* rather than *validating* or *empirical* evidence in Rychlak's sense (1959), for it involves belief in the Rorschach because of consensus among practitioners as to what the test means, its seeming consistency with clinical experience, and implied self-evident validity. The point is that though the Rorschach has not been validated unequivocally, the data from a Rorschach protocol have

a compelling effect on the behavior of the clinician vis-a-vis patient. Ironically, the clinician *qua* clinician is more apt to trust his "procedural" evidence over "empirical" evidence whenever they conflict, as is evident in the compartmentalization of daily clinical behavior from the findings of research, or articles of this type.

Construct validity in a test is usually attacked by hypothesizing some internal state, process or factorial composition of the test and finding its behavioral-relevance. In this connection there has been an increase in factorial studies of the Rorschach, a comprehensive review of which has recently been published by Murstein (1960). The results of 20 factorial studies reviewed by Murstein are far from conclusive since researchers disagree as to procedure and identity of factors isolated. As Murstein points out, some studies violate necessary assumptions of additivity and independence of variables; factors are rotated *ad libitum* until they substantiate the desired hypothesis; standard errors of factor loading are not known. Many of the psychometric weaknesses alluded to in this paper are cited as additional sources of inconsistent factorial results by Murstein. Procedural variations, i.e. type of Rorschach used, different scoring systems, different instructions to Ss, whole versus part analysis of tests, selection of experimental variables, combination and types of score treatment, indicate that "no two Rorschach matrices . . . have been identical". The net result of factorial efforts at establishing construct validity is pithily stated by Murstein: "from the point of research it (Rorschach) must be considered a psychometric sow's ear."

A third effort to give the Rorschach "silk-purse" validity has been to endow it with psychoanalytic qualities. Thus the idea of intrinsic construct validity in the Rorschach is given credibility by superimposing other theoretical frameworks in which

clinicians also believe. The Rorschach has been variously viewed as exemplifying Jungian psychology, classical and neo-classical Freudian theory. But this latter-day theorizing itself seems redundant inasmuch as the psychoanalytic theories themselves can lay little more claim for their construct validity than the self-same "procedural" evidence.

A SUMMATION

Viewing the Rorschach psychometrically, we find a plethora of mixed, sometimes contradictory psychometric methods operating, crude application of measurement principles, quasi-objective scoring, low reliability because of unknown triadic interactions among testee-tester, and groping, sometimes naive, sometimes sophisticated, for a valid rationale.

In defense of Hermann Rorschach, a bold innovator armed with a principle, he had described the present situation quite well in his introductory remarks to the second edition of *Psychodiagnostik* (1942):

"... At the outset it must be pointed out that all of the results are predominantly empirical. The questions which gave rise to the original experiments of this sort (1911) were of a different type from those which slowly developed as the work progressed. The conclusions drawn, therefore, are to be regarded more as observations than theoretical deductions. The theoretical foundation for the experiment is, for the most part, still quite incomplete..."

SOME TENTATIVE SOLUTIONS

✓ It is clear that the dilemma of the Rorschach is really the clinician's who is cognizant of its weaknesses. The author would like to propose some tentative solutions which will not reduce anxiety by lapsing into the tranquilizing formulation, "yet it works clinically", but by active research efforts. The following discussion is relevant to various issues raised in the article proper.

1. First and foremost is the ques-

tion: Shall we retain the Rorschach in good standing in our clinical armamentarium in view of its many weaknesses? In a word, *Yes!* It would make little sense to discard 2000 prior researches-whatever their limits-and some 30 years of clinical experience with the instrument. In all this formal and informal cumulative data there are admixtures of "true" and "error" variance. However, a drastic reappraisal about the premises of the Rorschach as well its operational limitations would seem to be in order.

2. Accept the Rorschach for what it is: a sheerly empirical technique. We need not camouflage this basic fact by superimposing equally dubious theoretical structures as an afterthought. The Rorschach as an adaptive task still has novel features, which can be utilized. The history of science is replete with many an empirical discovery which has proved useful despite its refractory nature from the point of view of theory, e.g. the early thermometer, aspirin, weather forecasting, shock therapy, hypnosis. The Rorschach technique may have to be ranked among these.

3. Shift our clinical focus to the observable here-now adaptive behavior of the subject in the Rorschach situation. This means that we have to forego luxuriating in the ephemeral domain of hypothetical constructs such as "projection", inferences about causality and genetic antecedents, unconscious forces and the like.

This approach has several possible merits that fit the Rorschach situation which we may operationally define as an interaction of person-test-environment. The author has become increasingly intrigued by the possible relevance of Helson's adaptation-level theory to the Rorschach situation.² This "theory" is in reality an

²As of this writing the author has obtained preliminary evidence of the operation of adaptation-levels in the Rorschach situation. (Block, 1961)

operational, quantifiable empirically-based set of relationships that account for the interaction of focal, background and contextual stimuli in a given situation. It is precisely the influence of these here-now adaptive influences vis-a-vis the patient's needs that loom as the hidden aspects of our Rorschachian iceberg. Since we can no longer cling to absolute concepts in testing in the Galtonian tradition, adaptation-level theory would represent the necessary infusion of relativistic concepts into the domain of projective testing.

Moreover by focusing on adaptive behavior, those very sets and response habits that were to-be-gotten-rid-of now become the-to-be-elicited. Stable sets would be indicative of stable adaptive traits in S. The clinician would henceforth need to explore in detail the nature of the sets operating in the Rorschach situation, particularly the confluence of conditions eliciting specific sets. What may be lost in terms of omniscience, that is, global sweeping generalizations about the patient which cannot be verified beyond personal clinical experience with all its pitfalls, may be gained in pertinence of data to behavioral potentialities of the patient.

4. Revise psychometric and scoring procedures used in the Rorschach. Certainly the overall Rorschach procedure may be strengthened by more explicit formulation of the task required of S, including certain caveats about sets or biases. But whatever limits are imposed by delineating instructions will probably be more than compensated for by increased reliability and trustworthiness of the obtained protocol.

The present scoring system, inordinately dependent upon the number of responses, would appear to be in need of overhaul toward meeting requirements of objectivity, additivity and scaling procedures involved. It might be desirable to construct a multiple-choice Rorschach that is prescored along more objective di-

mensions, and into which certain validity checks may be incorporated to detect dissimulation, idiosyncratic interpretations of instructions, etc.; as exist in the Minnesota Multiphasic Personality Inventory. These very same checks might be found to provide insight as to S's behavior in keeping with the focus on adaptation. Thus by restricting the bandwidth of information derived, we might obtain higher fidelity in the protocol.

The inquiry phase, if it is to be retained in the future, might very well be more structured to control its doubtful judgmental-introspective components. Baughman (1958) has already experimented with a paired comparison presentation during inquiry of a series of modifications of the properties of the original blot.

In keeping with the suggestion to concentrate on adaptive aspects in the Rorschach, greater use might be made of preferential sortings of the blots, and degrees of confidence in responses as factors influencing adaptation.

The above recommended solutions are just one concerned clinician's answer to an existential fact: the Rorschach is an empirical experiment continually modified by its creator to meet his clinical needs. Perhaps it is time for us to pick up the thread of experimentation left dangling by Rorschach's death, as he himself might have wished.

REFERENCES

- Anastasi, A. *Psychological testing*. New York: Macmillan, 1954.
- Baughman, E. E. A new method of Rorschach inquiry. *J. proj. Tech.*, 1958, 23, 381-389.
- Block, W. E. A preliminary study of adaptation-level theory as a framework for projective testing. Progress Report, N.I.M.H. Grant M-5118 (A) September 1961.
- Cronbach, L. J. Response sets and test validity. *Educ. psychol. Measmt.* 1946, 6, 475-494.
- Cronbach, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.* 1950, 10, 3-31.
- Cronbach, L. J. Statistical methods applied to Rorschach Scores: A review. *Psychol. Bull.*, 1949, 46, 393-429.

- Guilford, J. P. *Psychometric methods*. New York: McGraw-Hill, 1954.
- Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.* 1948, 55, 297-313.
- Masling, J. The influence of situational and interpersonal variables in projective testing. *Psychol. Bull.*, 1960, 57, 1, 65-85.
- Murstein, B. I. Factor analyses of the Rorschach. *J. consult. Psychol.* 1960, 24, 3, 262-275.
- Rabin, A. I. A contribution to the "Meaning" of Rorschach's Inkblots via the Semantic Differential. *J. consult. Psychol.*, 1959, 23, 4, 360-372.
- Rorschach, H. *Psychodiagnostics*. Berne: Hans Huber, 1942.
- Rosenzweig, S. Idiodynamics in personality theory with special reference to projective methods. *Psychol. Rev.*, 1951, 58, 213-223.
- Rychlak, J. F. Clinical Psychology and the Nature of Evidence. *Amer. Psych.*, 1959, 14, 10, 642-648.
- Schachtel, E. G. Subjective definitions of the Rorschach test situation and their effect on test performance. *Psychiat.*, 1945, 8, 419-448.

Received October 21, 1961.

The Rorschach Popular Response Among Hawaiian Schizophrenics

BERNARD L. BLOOM
Hawaii State Hospital

The concept of the popular response, first introduced by Rorschach, has remained one of the more durable aspects of the quantitative analysis. While there is still no specific general agreement regarding which responses are to be considered popular, Rorschach experts generally ascertain the extent to which a subject's percepts include those which are very commonly seen by some larger reference group. Ability to see what most people see is one aspect of psychological health. Rorschach (1942, p. 198) described the popular response as representing "... the share in the collective or common way of seeing or perceiving things." Klopfer and Kelley (1946, p. 216) suggest that "... not to use the most obvious concepts used by the great majority of other subjects may mean that the subject is not able to think along the lines of other people or that he is not willing to do so." Beck (1947, p. 16) indicates that the popular response projects "... the ability to participate in the common or popular thinking of the group, or the conformity of Ss thinking with that of the group." He states, furthermore (1952, p. 24), that "... the popular (P) response has been correctly evaluated as a measure of the ability to recognize the most common percepts of one's milieu. It points, therefore, to conformity in the thought content." Hertz (1938, p. 29) views the popular response of indicating "... the degree of stereotypy and commonality of thought processes." Bell (1948, p. 130) states that "... the popular responses are indicative of the amount of conformity of the individual with the average." Phillips and Smith (1953, p. 112) indicate that popular responses "... provide significant data about the

subject, particularly in relation to the socialization and the extent to which he has incorporated conventional patterns of behavior." Rapaport (1946, p. 315) suggests that "... the P responses have been considered to represent compliance with the thinking of the community—in other words, the capacity for thinking in conventional and stereotyped terms. Such a capacity is essential for balanced and realistic thinking."

As this brief review indicates, use of popular forms by a subject suggests that his perceptions are like those of the reference group, that he sees what they do and by implication that he tends to think the way they do. Insofar as the popular response is concerned, the Rorschach test is hardly projective. The areas which stimulate popular responses are not unstructured and vague; indeed, their relatively unambiguous structure makes the popular response possible. As Rapaport (1946, p. 315) points out, "... the areas to which popular responses are given represent a relatively clear-cut piece of reality which is so compelling that its 'meaning' is a matter of 'social agreement'. The responsiveness of a subject to these compelling areas on the ink blots thus becomes a measure of his sense for the 'obvious'."

The earliest definition of P was proposed by Rorschach as any response appearing in at least one protocol out of three. Viewing the P designation solely as based upon frequency of elicitation by a reference group seems both logical and straightforward and has found considerable support by research workers. The frequency of one in three has apparently seemed unduly demanding, however, and has been generally sup-

planted by one in six, e.g., by Hertz (1938) as well as by Ames and her co-workers (1952, 1954, 1959) and others. While there are several proposed arrays of *P* and while these show very considerable similarity with each other, the two most commonly used arrays were not derived by the foregoing method. The ten *P* of Klopfer and Kelley (1946) do not appear to have been developed with the use of any specified arithmetic criteria. The 21 *P* proposed by Beck et al. (1950) which include essentially all but two of the Klopfer *P* were based upon a complex arithmetic-inspectional analysis of the protocols obtained from 157 normal Chicago adults. In order to qualify as a *P* the area itself must have been chosen by at least 6% of the subjects, the response deemed popular must have been given by a minimum of 14% of the subjects and must have been at least three times as common as the next most common response to that area. Beck has accordingly lowered the one in six criterion slightly and has added two other criteria.

In view of the assumed clinical significance of *P* it might seem that number of *P* would be inversely related to extent of psychiatric disorder. Phillips and Smith (1953, p. 320) caution, however, that "the obtained *P* is remarkably impervious to the impact of psychopathology." There is considerable overlap between the distribution of number of *P* in normal and pathological groups, and research findings are not unequivocal in showing differences in average *P* when normals are compared with psychiatric patients and when various psychiatric groups are compared with each other. Molish (1951) compared *P* in Beck's Chicago sample of normal adults with *P* found in a sample of 60 neurotics and 80 psychotics and found that normals perceived significantly more *P* than either neurotics or psychotics but that neurotics did not differ from psychotics in the frequency of percep-

tion of *P*. In addition, he identified eight *P* responses which were given with significantly greater frequency by normals than by schizophrenics and four *P* responses which constituted in each case a higher proportion of all responses to these particular areas in the case of the normal subjects as contrasted with the schizophrenic group. These particular findings have not since been replicated, however. Warner (1951) found no significant differences between number of *P* in a group of outpatient schizophrenics and in a group of outpatient neurotics. He suggested that a decrease in *P* should be found among schizophrenics as chronicity and deterioration increases and proposed that length of hospitalization might be inversely related to number of *P* in a hospitalized schizophrenic sample.

The generality of the popular response is a question of considerable practical significance. The issue at hand is defining the reference group for which a particular response is popular. Beck makes it clear in his writings that no universal validity is claimed for his popular responses, that they apply only to North American adults. Hertz (1938), following her review of various proposed *P*, suggests six particular *P* which are probably universal, i.e., not affected by sex, age, race, or locality. Klopfer and Kelley (1946) distinguish between group populars characteristic of some particular age, racial, geographic, or cultural group, and universal populars which would theoretically be found in any and all such groups. Their list of ten populars purports to be applicable to an unlimited array of age and cultural groups. A study of Haitian protocols obtained by Bourguignon and Nett (1955) and a study of protocols obtained by Ewing and Stevenson (1953) from subjects in the provinces of Nova Scotia, Prince Edward Island and Newfoundland suggest that some of the Klopfer *P* have less than uni-

versal applicability. In reviewing this problem Hallowell points out (in Klopfer, 1956, pp. 509 ff) that "... our present empirical knowledge does not enable us to define the actual extent, or the transcultural limits, of 'universal' *P*'s. We can only determine these limits by assembling more data or group frequencies of all kinds." Thus, basic problems regarding the definition of *P* and the relationship of *P* to psychiatric status have been raised.

As part of a larger normative study of Rorschach responses, analysis is here presented of popular responses obtained in a sample of 101 hospitalized schizophrenics in Hawaii. These patients, nearly all of whom were born in Hawaii, range in age from 14 to 50 with a mean age of 31.5, and vary in educational level from 5th grade to college graduate with a mean of 10.5 grades completed. They comprise a varied racial and ethnic sample. Nearly half of the sample is Japanese; the remainder include (in decreasing frequency) Caucasians, part-Hawaiians, Filipinos, Chinese, Portuguese, Koreans, Hawaiians, and Puerto Ricans. Diagnostically, the group includes 51 paranoid schizophrenics, 17 catatonic schizophrenics, 14 chronic undifferentiated schizophrenics, eight schizo-affective schizophrenics, seven acute undifferentiated schizophrenics, three simple schizophrenics and one hebephrenic schizophrenic. The Rorschach test had been administered to these patients as part of the intake procedure with the testing usually completed within three weeks of admission. Follow-up information was generally available for a two-year period after discharge, but for never less than 18 months. These patients are geographically, culturally, racially, and diagnostically as different from the Beck normal Chicago sample as might be possible to obtain within the adult English-speaking world.

The purpose of this analysis is, first, to consider each Beck and Klop-

fer *P* as a hypothesis to be evaluated in the present data. To the extent that the *P* also emerges in the present sample, the base for that particular *P* is considerably broadened. The second purpose is to provide additional data regarding the relationship of *P* responses to degree of psychopathology.

RESULTS

Comparison of the Chicago normal adult sample and the Hawaii schizophrenic sample as regards frequency of the Beck and Klopfer *P* responses is found in Table I. Information regarding the Chicago sample has been adapted from Beck et al. (1950, pp. 280-282). Calculation of *P* by the one protocol out of six criterion in the Hawaii schizophrenic sample yields six of the traditional *P*: the bat, butterfly, or moth for Card I, W; human figures for Card III, D1; bat, butterfly, or moth for Card V, W; animal hide, pelt, skin, or rug for Card VI, W or D1; animal for Card VIII, D1; and the crab, lobster, or spider for Card X, D1. Calculation of *P* by Beck's criteria adds two to the list: human figures on Card II, W; and dog or bear on Card II, D1. While the bow-tie or ribbon response for Card III, D3 is not *P*, in combination with the butterfly response as Klopfer suggests, the *P* designation is justified. By either criterion, then, nine *P* emerge of which all are in the Beck list and all but the human figures on Card II, W are on the Klopfer list. Examination of the protocols does not yield any *P* not on either Beck's or Klopfer's list.

The Ss in the Chicago group were considerably more responsive than were the Hawaii Ss. While this may be a function of the differences in psychiatric status, previous studies of Rorschach protocols given by schizophrenic patients have shown them to be nearly as responsive as normal subjects. Thus, the difference in responsivity between the present groups is at least partly a cultural phenom-

TABLE I Comparison of Beck and Klopfer P in Chicago and Hawaii Samples

			Chicago Sample (N=157)			Hawaii Sample (N=101)		
			R	P	P/R%	R	P	P/R%
Card I, W Card I, D3 or D4	(B)	Beck (B) and Klopfer (K) Popular Responses						
	(B)	Bat, butterfly or moth	228	114	51**	123	36	29
	(B)	Figure of a person	85	52	61	28	13	46
Card II, W Card II, D1 Card II, D3	(B)	Human figures	61	34	56**	54	15	28
	(B)	Dog or bear	147	83	56	36	14	39
	(B)	Butterfly or moth	60	22	37	12	2	17
Card III, D1 Card III, D3 Card III, D3	(B)	Two humans	144	93	65	80	47	59
	(B)	Bow-tie or ribbon	60	28	47**	29	6	21
	(B)	Bow-tie, hair ribbon or butterfly	60	46	72	29	17	59
Card IV, W Card V, W Card V, D1	(B)	Animal hide, pelt, skin or rug	163	36	22	77	12	16
	(B)	Bat, butterfly or moth	207	139	67*	109	60	55
	(B)	Human or animal leg	34	30	88	2	1	50
Card VI, W or D1 Card VII, D1	(B)	Animal hide, pelt, skin or rug	155	90	58**	72	28	39
	(B)	Human heads or faces	60	27	45	9	2	22
Card VIII, D1 Card VIII, D3 Card VIII, D4	(B)	Animal, bears, cubs, mice, etc.	195	113	58	64	45	70
	(B)	Skeletal form or ribs	29	25	86**	6	2	33
	(B)	Tree or bush	60	35	58	13	4	31
Card IX, D3 Card IX, D4	(B)	Human figure	61	57	93**	15	6	40
	(B)	Human head or face	44	42	95**	22	12	55
Card X, D1 Card X, D2 Card X, D5	(B)	Crab, lobster or spider	107	64	60	31	19	61
	(B)	Dog	66	34	52	12	4	33
Card X, D12 Card X, D4	(B)	Rabbit's head	46	25	54	12	9	75
	(B)	Sheep	46	24	52**	10	0	0
	(B)	Caterpillar, snake or worm	not given			11	7	64

**Significantly greater than the Hawaii sample at the 0.01 level.

*Significantly greater than the Hawaii sample at the 0.05 level.

enon. In view of the greater general responsivity of the Chicago sample compared with the Hawaii sample, one would anticipate that a higher proportion of Chicago Ss than Hawaii Ss would give any particular *P*. Examination of the data supports this proposition. When each *P* response is viewed in terms of the proportion of Ss giving it, the proportion is consistently higher in the Chicago sample than in the Hawaii sample. In the case of 17 of the 22 *P* (comparable data not being available regarding one Klopfer *P*) these differences are significant at the 0.01 level or better. Viewing each *P* response as a proportion of the total number of responses given to a particular area, there are fewer large differences, but again every significant difference favors the Chicago sample. Since this proportion is not affected by the differential responsivity in the two groups, these differences likely reflect psychiatric rather than cultural factors, i.e., greater perceptual variability in the schizophrenic group.

The present results confirm ten of Molish's (1951) 12 significant findings already referred to regarding certain *P* responses given significantly more often by normals than by schizophrenics. The exceptions are the human figures on Card III, D1 which in his sample was given by a significantly smaller proportion of schizophrenics than normals and the crab response on Card X, D1 which constituted a significantly higher proportion of all the responses to that area in the normal group than in the schizophrenic group. With respect to this latter difference, Bloom (1959) has shown that Hawaii ecology might be expected to result in a lowered perceptual threshold to this particular animal form.

Regarding the present confirmation of Molish's findings, it should be pointed out that comparison of the frequency of *P* responses in any Rorschach area between the Chicago normal sample and the Hawaii

schizophrenic sample generally favored the Chicago sample. Of the 44 statistical comparisons made in the present study, nearly 60% yielded differences significant at the 0.01 level or better. One would accordingly expect that seven of Molish's 12 specific findings would be confirmed on a chance basis alone. Confirmation of ten of his specific findings is not significantly different from this chance expectation. While the present data are consistent with his results, they cannot be used as an unequivocal test of them.

Of the nine *P* responses established in the Hawaii sample, the average number of *P* among male patients is 3.2 and among female patients is 2.3. This difference is significant at the 0.02 level ($t = 2.26$). Analysis of sex differences in each *P* response within the Hawaii sample revealed that male patients gave the animal hide response to Card VI, W or D1 and the animal to Card VIII, D1 significantly more frequently than female patients. The animal hide response was given by 43% of the male patients and only 18% of the female patients ($t = 2.8$). The animal response on Card VIII was given by 62% of the male patients and 34% of the female patients ($t = 2.8$). Sex differences in *P* have not been adequately investigated and the present findings require verification and elaboration.

At least four suggestions have appeared in the literature for lists of "universal" *P* responses. As was already indicated, the ten Klopfer *P* responses are viewed as universal by their authors. Of these responses, two fail to achieve *P* status in the present protocols. In both cases (Card X, D4 and Card X, D5) the areas themselves do not elicit enough responses to be eligible for consideration as *P*. Hertz (1938) suggests six universal *P* of which five appear in the present list. The exception is the animal skin for Card IV, W. Griffin (1957) lists six universals and all appear as *P* in the present sample. Hal-

lowell (in Klopfer, 1956) suggests four universals of which three appear in the present sample. The exception in this case is the rabbit head on Card X, D5. Comparing the four suggested lists of universal *P* reveals that only one response (animal, Card VIII, D1) is common to all lists. This response is also *P* in the Hawaii schizophrenic sample and thus appears to be the one unequivocal universal *P* reported in the literature thus far.

The *P* which appear on any three of the four lists (bat on Card I, W; animals on Card II, D1; human figures on Card III, D1; winged creature on Card V, W; and animal skin on Card VI, W or D1) are all *P* in the present sample and thus may warrant being designated as near-universal *P*.

In order to study the relationship between number of *P* responses and severity of psychopathology, the sample of 101 patients was subdivided into four groups according to their psychiatric history following their first admission to the hospital. Group I (five males and 14 females) consisted of those patients who were first admissions to the hospital, whose length of hospitalization was less than six months and who did not return to the hospital during the follow-up period. Group II (8 males and 7 females) were likewise first admissions to the hospital who did not return to the hospital during the follow-up period, but their length of hospitalization varied between six months and one year. Group III (15 males and 11 females) consisted of all first admission patients who were re-hospitalized following their discharge and those first admission patients whose length of hospitalization was in excess of one year regardless of subsequent treatment history. Group IV (9 males and 32 females) consisted of all patients with more than two admissions to the hospital and included those patients who had not been discharged following their second admission. It is felt that these group-

ings reflect increasing severity and chronicity of psychopathology.

Tabulating *P* only among the nine responses shown to be adequately substantiated in the present sample, the relationship between number of *P* and degree of psychopathology is shown in Figure 1. For the total group, the mean *P* is 2.7 and the standard error is 0.2. Sex differences in number of *P*, already referred to,

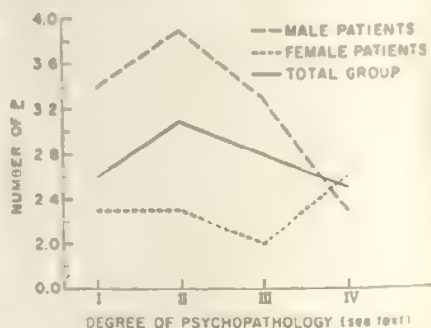


FIG. 1.
Relation Between Degree of Psychopathology and Number of *P* Responses.

are seen from this analysis to characterize all but the most chronic group. While there is relatively little difference within the female group as regards the relationship of *P* to degree of psychopathology, in the case of the male patients the relationship is quite striking. While the number of cases in each group is rather small, these results do suggest that a significant non-linear relationship may exist between degree of psychopathology and its manifestation in the Rorschach *P* response.

In view of the fact that seven of the nine *P* in the Hawaii sample are actually the most commonly seen *P* in the Beck normal sample, it may be that within the nine *P* it would be fruitful to take into account their relative popularity. Taking this into account would make it possible to distinguish a person giving a number of very commonly seen *P* responses from another individual giving an

equal number of less frequently seen *P*. Analysis of the nine *P* in the present study reveals that the most commonly seen *P* (Card V, W-bat, butterfly or moth) is more than four times as frequently mentioned as the least commonly seen *P* (Card II, D1-dog or bear). In order to examine the usefulness of a distinction within *P* responses according to relative popularity, an arbitrary weighting was assigned to each *P* so that weights between one and five were given to the nine *P* in direct relationship to how commonly each *P* was seen in the total group. A total score was then calculated for each patient which was based on the weighted sum of *P* rather than on the raw number of *P*. In Figure 2, the weighted *P* scores are shown as a function of psychopatho-

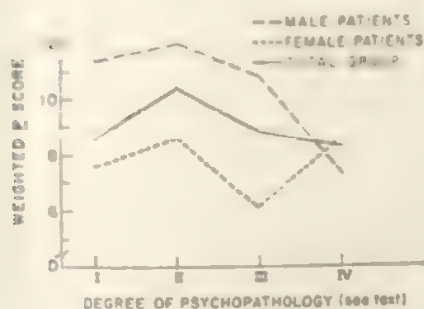


FIG. 2.

Relation Between Degree of Psychopathology and Weighted *P* Responses.

logy. Regarding the weighted *P* scores, the mean for the total group is 8.8 and the standard error is 0.5. The weighted *P* scores thus tend to reveal in greater magnification the changes with increasing chronicity and suggest as an additional observation that changes in the nature of the *P* responses with increasing psychopathology occur in a different sequence in the two sexes. The hypothesis may be advanced that changes in overt behavior with increasing psychopathology may also occur in a different sequence when male and fe-

male schizophrenic patients are contrasted.

DISCUSSION

Determination of *P* in a sample of Hawaiian schizophrenics has resulted in an array of nine responses. All of these responses have been previously suggested as *P* in a sample of clinically normal Chicago adults. The normative basis of these responses has accordingly been broadened both with respect to psychiatric status and racial/ethnic background. These nine *P* are not a random sample of the 23 Beck and Klopfer *P*. Rather, seven are the most commonly seen *P* in the Beck normal sample. Degree of popularity thus may be a relevant variable to consider in evaluating those *P* responses given by a particular subject. *P* might quite appropriately be viewed as one extreme in a frequency continuum of responses to any location. With appropriate normative data, every response on the Rorschach test could be given a value to indicate how popular it is. The so-called Original response would represent one end of the continuum and a response most commonly seen in the reference group might represent the other end of the same continuum. An approach of this kind was suggested by Meltzer (1954) and more recently a similar approach was attempted by Griffin (1957). These methods would substitute a psychometric scaling procedure for the traditional *P* concept. Availability of such a scale would significantly simplify cross-cultural comparisons as well as other group comparisons and would open the way to examine such measures as the average popularity of responses in a particular protocol or the variability in popularity scores within or between protocols. Such a scale, for example, would also permit the identification and investigation of unpopular responses.

There is no simple influence of ecologic factors on the particular *P* responses given in the Hawaii schizo-

phrenic sample. On Card I, W, for example, 25 of the 36 *P* responses involve some reference to "bats". Of the 25 patients who gave these responses, 18 were born in Hawaii. Yet there are no bats in Hawaii. Butterflies and moths are extremely common, however, but they are not nearly as commonly seen in this inkblot area as bats. The mongoose, for another example, is a distressingly common rodent in Hawaii. Yet to Card VIII, D1, perception of mice or rats is three times as frequent as that of mongooses which that particular area resembles about as well as it does other rodents. Colored areas throughout the inkblots provoke frequent responses of flowers or coral formations, however, but no adequate normative data are easily available to evaluate how much more frequently these responses occur in the Hawaii sample than in another reference group.

As regards the relationship of severity of pathology and frequency of *P* responses, the results support previous findings relating Rorschach scoring categories to degree of psychopathology. The acutely disturbed patient who provides a disrupted Rorschach protocol, including few *P* responses (as in the case of our Group I patients) actually shows a rapid and lasting clinical improvement. Patients whose perceptual processes are not as disturbed frequently show less rapid clinical improvement (as in the case of many of our Group II patients). With still greater clinical disability (as in our Group III and then Group IV patients) perceptual disorganization again increases. Thus it may be both the patients with the best and poorest prognosis who exhibit the greatest perceptual disruption.

SUMMARY

A normative study of the *P* response has been undertaken in a sample of 101 hospitalized Hawaiian schizophrenics in order to investigate some socio-cultural and psychiatric

dimensions of the *P* designation and to explore further the *P* response as an index of psychopathology. No new *P* responses were found in this sample. Rather nine responses, all previously suggested as *P* on the basis of normal samples were delineated. Of these responses, seven are the most commonly seen in the normal reference group. The suggestion for a psychometric scale of popularity to replace the traditional *P* vs. non-*P* designation was advanced because of the relevance of the degree of popularity. Sex differences were found with male patients perceiving significantly more *P* than female patients. Results of the investigation suggest that there may be a non-linear relationship between degree of psychopathology and frequency of perception of *P*. Most *P* were perceived by those male patients with a moderately good clinical outcome. Fewest *P* were perceived by both patients with the best and the poorest clinical history.

REFERENCES

- Ames, Louise B. et al. *Child Rorschach responses*. New York: Harper, 1952.
- Ames, Louise B. et al. *Rorschach responses in old age*. New York: Harper, 1954.
- Ames, Louise B. et al. *Adolescent Rorschach responses*. New York: Harper, 1959.
- Beck, S. J. *Rorschach's test*. Vol. 2. *A variety of personality pictures*. New York: Grune & Stratton, 1947.
- Beck, S. J. *Rorschach's test*. Vol. 3. *Advances in interpretation*. New York: Grune & Stratton, 1952.
- Beck, S. J. et al. The normal personality as projected in the Rorschach test. *J. Psychol.*, 1950, 30, 241-298.
- Bell, J. E. *Projective techniques*. New York: Longmans, Green, 1948.
- Bloom, B. L. Ecologic factors on the WAIS picture completion test. *J. consult. Psychol.*, 1959, 23, 375.
- Bourguignon, Erika E. & Nett, Emily W. Rorschach populars in a sample of Haitian protocols. *J. proj. Tech.*, 1955, 19, 117-124.
- Ewing, R. M. & Stevenson, Margaret. A report of research on Rorschach "populars". *Bull. Marit. psychol. Assn.*, 1953 (April), 15-18.
- Griffin, Dorothy P. Psychometric scales for the Rorschach popular response. *J. clin. Psychol.*, 1957, 13, 283-287.
- Hertz, Marguerite R. The "popular" re-

- sponse factor in the Rorschach scoring. *J. Psychol.*, 1938, 6, 3-31.
- Klopfer, B. & Kelley, D. M. *The Rorschach technique*. New York: World Book Co., 1946.
- Klopfer, B. et al. *Developments in the Rorschach Technique*. Vol. II, *Fields of application*. New York: World Book Co., 1956.
- Meltzer, H. Personality differences among stuttering children as indicated by the Rorschach test. *Amer. J. Orthopsychiat.*, 1934, 4, 262-282.
- Molish, H. B. The popular response in Rorschach records of normals, neurotics and schizophrenics. *Amer. J. Orthopsychiat.*, 1951, 21, 523-531.
- Phillips, L. & Smith, J. G. *Rorschach interpretation: advanced technique*. New York: Grune and Stratton, 1953.
- Rapaport, D. *Diagnostic psychological testing*. Vol II. Chicago: Year Book Publishers, Inc., 1946.
- Rorschach, H. *Psychodiagnostics*. (4th ed.) New York: Grune & Stratton, 1942.
- Warner, S. J. An evaluation of the validity of Rorschach popular responses as differentiae of ambulatory schizophrenia. *J. proj. Tech.*, 1951, 15, 268-275.

Received November 22, 1961.

The Validation of Projective Tests¹

RICHARD H. DANA
West Virginia University

Reviewers of projective tests no longer find them "promising" (Eysenck, 1959; Shaffer, 1959). This criticism may reflect the absence of an acceptable frame of reference within which the research literature may be evaluated. Three postulates about validity, taken from one such framework (Dana, 1960a), are considered here. A. Projective tests provide data from which there is a high probability of making accurate predictions about human beings. B. Projective tests provide behavior samples which are positively correlated with external observations of the same behavior. C. Projective tests elicit data that are positively related to theoretically derived constructs which are descriptive of human beings.

VALID FOR WHAT?

No test can claim equal utility for many unrelated measurement purposes. Projective tests are used as measures of personality characteristics. Predictions are made from test data to overt behaviors and to specific variables in the present (*i.e.*, psychiatric diagnosis) or future (*i.e.*, outcome of treatment or hospitalization). Such diversity of claims was partially responsible for the formal definition by an APA committee of kinds of validity (1954), content, predictive, concurrent, construct, and the later addition of a fifth, congruent (Thorn-dike & Hagen, 1955). The question, "valid for what purpose?", is infrequently asked in discussions of validity or in test reviews (exceptions are Shneidman, 1959; Dana, 1959c). Nonetheless, tests are more efficient for some measurement objectives than for others.

POSTULATE A: PREDICTIVE VALIDITY

Since predictive, concurrent, and congruent validity differ primarily with respect to time and nature of criterion, they will be considered together. Concurrent validity refers to prediction of present status variables; *e.g.*, diagnosis, while predictive validity is concerned with future performance; *e.g.*, outcome of treatment. Congruent validity refers to the relationship between the test and other test measures of the same variable.

Windle's review of predictive and concurrent validity studies found little merit in projective tests (1952). Windle, however, formulated criteria for future studies. Validation attempts must specify experimental conditions, use homogeneous subjects, apply relevant statistics, and be cross validated. In spite of more sophisticated experimental designs, some test variables meet the criterion of cross validation (*e.g.*, Dana, 1956; Powers & Hamlin, 1955; Taulbee & Sisson, 1954), while others do not (*e.g.*, Auld & Eron, 1953).

Individual scores or "signs" have been used historically for predictive validation. However, generalizability of "signs" may be limited by inadequate reliability (test score and scorer) and situational artifacts. Global judgments in which clinical skill overshadows the use of particular scores may be preferable (Zamansky & Goldman, 1960). Nonetheless, part of the problem may lie in the attempt to relate a test score with an intrinsically unrelated criterion in the absence of an explicit rationale. Predictions which are hypotheses may be confirmed or disconfirmed. Predictions which are random speculations have no legitimate place in test validation procedure.

¹ Adapted from a paper read at the 16th International Congress of Psychology, Bonn, Germany, August, 1960.

Two consequences of this use of "signs" are found in infrequent successful cross validation and substitution of simpler, shorter, more economical, non-projective tests. We do need to know the kinds, or general classes, of variables which will survive cross validation. However, before this knowledge is available, other tests, such as those used by Zubin in prognosis studies, may have supplanted projective instruments (Zubin *et al.*, 1961). Loevinger (1957) has remarked that predictive validation seems to involve "egg-boiling decisions."

Although the literature abounds in congruent validity studies, the value of many such reports can be questioned. First, "buckshot" attempts to find empirical correlates may yield only consternation (Dana, 1960b). Again, rational thought and hypotheses appear mandatory. Secondly, when specific hypotheses are investigated via correspondence in various projective tests, limited support is found (Walker, 1951). Third, the degree of similarity in test results is roughly related to the level of measurement provided by the specific test or scoring variable. Thus, Rorschach-TAT or Machover-TAT comparisons (Gallesse & Spoerl, 1954; Shatin, 1955) are more striking than are Rorschach-MMPI, TAT-SCT, or TAT-MMPI comparisons (Carr, 1954; Child *et al.*, 1956; Dana & Mueller, 1961; Rosen, 1952).

POSTULATE B: CONTENT VALIDITY

Content validity refers to the adequacy of the test or test variable as a sample of present behavior. Since both test and criterion are, in effect, samples from the same universe, the correspondence may not be great. The criterion should be stable, relevant, unbiased, available, easily defined, and prototypical. Such criteria for content validity studies have not been forthcoming.

One example will illustrate the complexity of relationships between

projective test data and content criteria. Nine studies used different projective tests; *e.g.*, Rorschach, TAT, MAPS, to measure hostility-aggression. Each study used different test variables and different criterion measures thus rendering comparability between studies using the same test impossible. No relationship (Gluck, 1955; Rader, 1957), or, at best, a slight, positive relationship between test variable and hostility-aggression criterion was found (Finney, 1955; Wolf, 1957). This relationship is affected by the number of projective tests used (Smith & Coleman, 1956), by social class variables (Mussen & Naylor, 1954), by maternal attitudes toward aggression (Lesser, 1957), and by impinging, inhibitory personality processes (Mussen & Naylor, 1954; Purcell, 1956). A more general conclusion is that behaviors which are least subject to disguise may be best predicted (Fisher & Morton, 1957).

POSTULATE C: CONSTRUCT VALIDITY

Construct validity has been caricatured as responsible for "a non-empirical, nonscientific approach to the study of behavior" (Bechtoldt, 1959, p. 628) whose technical problems are "similar to measuring a floating cloud with a rubber band in a shifting wind" (Shneidman, 1959, p. 261). Concern with construct validation is the result of behavioral or personality phenomena which are too complex for classical methodology. When such complexity occurs, man resorts to common sense, judgment, and intuition. Rychlak (1959) has considered two kinds of evidence: validating, in which there are observable consequences, and procedural, in which consistency, intelligibility, and communal common sense combine into a lay approximation of construct validity. More technically, concern with construct validity occurs whenever there is a nonoperationally defined concept (Cronbach & Meehl, 1955), a hypothetical or postulated underlying characteristic in the ab-

sence of a definite criterion (APA, 1954).

Construct validation is established by several lines of converging evidence (Loevinger, 1957) according to criteria of fidelity, meaning, and prediction. Fidelity means that the construct is reflected in test item content. Meaning is given by the completeness of the network of established empirical relationships in which the construct occurs. Predictions based on available knowledge of the construct must be consistent with external correlations.

The validation of Rorschach's human movement construct, M, was attempted by Singer and his coworkers (1956). Rorschach M involves the ability to perceive movement in static blots. The literature on Rorschach M was reviewed in order to provide a definitional context. A series of predictions were made between specified M components and a variety of test and task variables. Then, a network of correlations (empirical relationships) between fantasy and control of motility and impulsive behavior was generated. The magnitude of the obtained relationships confirmed the predictions. Thus, the criteria of fidelity, meaning, and prediction are met.

Singer subjected the matrix of relationships to a factor analysis and discovered four centroid factors. Inspection of factor loadings suggested that "the original formulation was overly simplified" (Singer *et. al.*, 1957, p. 382). This example clearly indicates the present status of Rorschach M. Unfortunately, however, there are only a limited number of such studies.

Projective tests were not intended primarily as psychodiagnostic instruments, nor as predictors of discrete, outcome variables. They were designed to facilitate the conceptualization of an implicit network of hypothetical relationships among personality characteristics of one individual.

Predictions are made from this network to the variety of relevant behaviors in which the clinician may have interest. However, validation of such predictions logically comes *after* construct validation. When content validity, for example, precedes construct validity, then the typical finding is an attenuated relationship demonstrable under certain conditions where adequate control permits specification of other relevant variables and relationships. The nine hostility-aggression studies previously discussed exemplify this circumstance.

Validation of projective tests should focus on the construct status of major test variables and concepts. In this way it will also be possible to investigate the process of clinical judgment. The facile clinician, in the absence of a panoply of construct validity studies, is forced to systematize relationships among test variables and all of his available knowledge of external correlates. This process occurs at various levels of awareness and with varying degrees of labeling the relevant cues. Since clinicians differ in knowledge, confidence, and awareness, a normal distribution of clinical skill is not unexpected (Hunt *et. al.*, 1954). Nor is it surprising that the range of clinical skill appreciably diminishes as the conditions of judgment are made explicit and uniform (Hunt, 1959). Contextual networks of known relationships among personality characteristics can provide impetus to studies of clinical judgment.

In addition, construct validity offers data from a systematic method as an immediate substitute for personality theory. There have been few studies testing deductions from personality theory in recent years. The preoccupation with part-process under laboratory conditions does not seem likely to result in the development of personality theory. Construct validation studies, using projective test variables interrelated with theoretically derived predictions, can be-

come a basis for future personality theory.

SUMMARY AND CONCLUSIONS

Three postulates involving empirical, content, and construct validation were discussed. Empirical validities; i.e., concurrent, predictive, congruent, appear to be relatively minor concerns of projective tests. Content validity, in the absence of suitable external criteria, becomes a potent method for generating disagreement and confusion. Construct validity appears to be the appropriate direction for projective test validation.

Construct validity involves the convergence of several lines of evidence in accord with criteria of fidelity, meaning, and prediction. There are three major values to construct validation. (a) Projective test interpretation can be based eventually on the personality theory contained in a large number of construct validation studies. (b) Construct validity may be seen as a formal model for clinical judgment; in other words, a method of making clinical concepts more explicit. (c) Validation of specific predictions, empirical validity, logically comes *after* construct validity is established. Prediction is then from a network of known empirical relationships rather than from a set of test scores. Construct validity is a substitute for piecemeal validation research which may lead to increments of knowledge well beyond the particular constructs investigated.

REFERENCES

- American Psychological Assn., Committee on Psychological Tests. *Technical Recommendations for Psychological Tests and Diagnostic Techniques*. Washington, D.C.: APA, 1954.
- Auld, F., Jr., & Eron, L. D. The use of Rorschach scores to predict whether patients will continue psychotherapy. *J. consult. Psychol.*, 1958, 17, 104-109.
- Bechtoldt, H. P. Construct validity: a critique. *Amer. Psychol.*, 1959, 14, 619-630.
- Carr, A. C. Intra-individual consistency in responses to tests of varying degrees of ambiguity. *J. consult. Psychol.*, 1954, 18, 251-258.
- Child, I. L., Frank, Kitty F., & Storm, T. Self-ratings and TAT: Their relations to each other and to childhood background. *J. Pers.*, 1956, 25, 96-114.
- Cronbach, L. J., & Meehl, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, 52, 281-303.
- Dana, R. H. A proposed theoretical framework for projective test literature. Unpublished manuscript, University of Nevada, 1960. (a)
- Dana, R. H. Objective TAT scores and personality characteristics: Perceptual Organization. *Percept. mot. Skills*, 1960, 10, 154. (b)
- Dana, R. H. The Picture-Frustration Test. In Buros, O. (Ed.) *5th Mental Measurements Yearbook*. Highland Park: Gryphon, 1959. (a)
- Dana, R. H. Cross validation of objective TAT scoring. *J. consult. Psychol.*, 1956, 20, 33-37.
- Dana, R. H., & Mueller, D. Congruent validation of a TAT score: Perceptual Organization. *J. Psychol. Stud.*, 1961, in press.
- Eysenck, H. J. The Rorschach Test. In Buros, O. (Ed.) *5th Mental Measurements Yearbook*. Highland Park: Gryphon, 1959.
- Finney, B. C. Rorschach test correlates of assaultive behavior. *J. proj. Tech.*, 1955, 19, 6-17.
- Fisher, S., & Morton, R. Levels of prediction from the TAT. *J. consult. Psychol.*, 1957, 21, 115-120.
- Gallese, A. J., & Spoerl, Dorothy T. A comparison of Machover and Thematic Apperception Test interpretation. *J. soc. Psychol.*, 1954, 40, 73-77.
- Gluck, M. R. The relationship between hostility in the TAT and behavioral hostility. *J. proj. Tech.*, 1955, 19, 21-26.
- Hunt, W. A. An actuarial approach to clinical judgment. In Bass, B., & Berg, I. (Eds.) *Objective Approaches to Personality Assessment*. Princeton: Van Nostrand, 1959.
- Hunt, W. A., Arnhoff, F. N., & Cotton, J. W. Reliability, chance, and fantasy in interjudge agreement among clinicians. *J. clin. Psychol.*, 1954, 10, 294-296.
- Lesser, G. S. The relationship between overt and fantasy aggression as a function of maternal response to aggression. *J. abnorm. soc. Psychol.*, 1957, 55, 218-221.
- Loevinger, Jane. Objective tests as instruments of psychological theory. *Psychol. Rep.*, 1957, 3, 635-695.
- Mussen, P. H., & Naylor, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.
- Powers, W. T., & Hamlin, R. M. The relationship between diagnostic category and deviant verbalizations on the Rorschach. *J. consult. Psychol.*, 1955, 14, 120-124.
- Purcell, K. The TAT and antisocial behavior. *J. consult. Psychol.*, 1956, 20, 449-456.
- Rader, G. E. The prediction of overt aggression.

- sive verbal behavior from Rorschach content. *J. proj. Tech.*, 1957, 21, 294-306.
- Rosen, E. MMPI and Rorschach correlates of the Rorschach white space response. *J. clin. Psychol.*, 1952, 8, 283-288.
- Rychlak, J. F. Clinical psychology and the nature of evidence. *Amer. Psychol.*, 1959, 14, 642-648.
- Shaffer, L. F. The Rorschach Test. In Buros, O. (Ed.) *5th Mental Measurements Yearbook*. Highland Park: Gryphon, 1959.
- Shatin, L. Relationships between the Rorschach test and the Thematic Apperception Test. *J. proj. Tech.*, 1955, 19, 317-331.
- Shneidman, E. S. Suggestions for the delineation of validation studies. *J. proj. Tech.*, 1959, 23, 259-263.
- Singer, J. L., Wilensky, H., & McCraven, Vivian. Delaying capacity, fantasy, and planning ability. *J. consult. Psychol.*, 1956, 20, 375-383.
- Smith, J. R., & Coleman, J. C. The relationship between manifestations of hostility in projective tests and overt behavior. *J. proj. Tech.*, 1956, 20, 326-334.
- Taulbee, E. S., & Sisson, B. D. Rorschach pattern analysis in schizophrenia: a cross validation study. *J. clin. Psychol.*, 1954, 10, 80-82.
- Thorndike, R. L., & Hagen, Elizabeth. *Measurement and Evaluation in Psychology and Education*. New York: Wiley, 1955.
- Walker, R. A. A comparison of clinical manifestations of hostility with Rorschach and MAPS tests performances. *J. proj. Tech.*, 1951, 15, 444-460.
- Windle, C. Psychological tests in psychopathological prognosis. *Psychol. Bull.*, 1952, 49, 451-483.
- Wolf, I. Hostile acting out and Rorschach test content. *J. proj. Tech.*, 1957, 21, 414-419.
- Zamansky, H. S., & Goldman, A. E. A comparison of two methods of analysing Rorschach data in assessing therapeutic change. *J. proj. Tech.*, 1960, 24, 75-83.
- Zubin, J., Sutton, S., Salzinger, K., Salzinger, Suzanne, Burdock, E., & Peretz, D. A biometric approach to prognosis in schizophrenia. In Hoch, P., & Zubin, J. (Eds.), *Comparative Epidemiology of the Mental Disorders*. New York: Grune & Stratton, 1961.

Received October 6, 1961.

Rorschach Developmental Level and the MMPI as Measures of Severity of Psychological Disturbance¹

MARVIN R. GOLDFRIED
University of Rochester

An important dimension within the general area of psychopathology has been the intensity or severity of the psychological disturbance. One approach to severity of pathology views the problem within a framework of developmental theory. Werner (1948), for example, has formulated a concept of personality organization by ordering an individual's level of functioning along a developmental continuum. Werner maintains that cognitive development progresses according to the principle of *orthogenesis* (Werner, 1957); i.e., it proceeds from a less differentiated to a more articulated mode of functioning. This increasing differentiation and hierarchical integration proceeds along several polar dimensions: syncretic-discrete, diffuse-articulated, indefinite-definite, rigid-flexible, and labile-stable. Werner further maintains that pathologically disturbed individuals are in a state of regression, and consequently are functioning at levels of cognitive organization similar to (though not identical with) those of children. In general, the greater the severity of disturbance, the more the regression, and consequently the lower the level of development.

The identification and investigation of developmental levels of personality organization has been approached by several researchers by means of the Rorschach. Friedman (1953) developed a Rorschach scor-

ing system to parallel Werner's genetic theory of cognition. This scoring system differs from the conventional Rorschach scoring in that only the structural and organizational aspects of the percept are used. Consequently, Friedman has used location scores only, classifying the percepts according to the level of diffuseness, articulation, integration, etc. For example, a genetically low response might consist of an amorphous whole (W_a) of "black paint" to card I, while a genetically high response would be a well-articulated and integrated whole ($W++$) of "a giant sitting on a stump" to card IV. The reliability of this scoring system showed agreement between judges ranging from 89.7 per cent to 95.5 per cent. Validation studies of this method of scoring have consistently yielded positive results. Hemmendinger (1953) found that genetically low scores decreased and genetically high scores increased between the ages of 3 and 10; Siegel (1953) tested and confirmed the hypothesis that catatonic and hebephrenic schizophrenics are more regressed than paranoid schizophrenics, finding the latter group to give percepts of a higher developmental level. Becker (1956), and even more recently, Fine & Zimet (1959), have shown that "reactive" schizophrenics have a higher developmental level than "process" schizophrenics. The results of these and other studies have demonstrated that this scoring system represents a reliable method of assessing developmental level of perceptual organization and is an accurate indicant of severity of psychological disturbance.

Characteristic of this developmental approach to psychopathology is

¹ The data for this study were collected while the author was at the Veterans Administration Hospital, Palo Alto, California. A portion of this paper was presented at East. Psychol. Ass'n, Philadelphia, April, 1961. The author would like to thank Marvin J. Feldman and Philip A. Goldberg for their helpful comments.

the fact that it is theoretical in nature. Theory determines the predictions that are made and has even entered into the construction of the measuring instrument itself. By way of comparison, the Minnesota Multiphasic Personality Inventory (MMPI) has utilized very little theory in the investigation and diagnosis of severity of pathology. While there undoubtedly was some implicit theoretical influence in the formulation of the MMPI items, the clinical scales themselves were derived solely on their ability to differentiate predetermined psychiatric groups.

With the continued use of the MMPI, the emphasis has moved from the use of single elevated scores to an analysis of the configuration or pattern of scores in the interpretation of a record (Gough, 1946; Schmidt, 1945; Sullivan & Welsh, 1952). The pattern analysis approach has added greatly to the usefulness of the MMPI in differential diagnosis and in the assessment of severity of pathology. An early, though crude attempt at formalizing the pattern approach to the MMPI was attempted by Ruesch & Bowman (1945). By taking the sum of scores on Hs, D, and Hy, they arrived at an overall neurotic (N) score. A psychotic (P) score was derived by adding the scores on Pa, Pt, and Sc. In an attempt to demonstrate the utility of employing the total pattern of scores, Meehl (1946) sorted 147 "abnormal" (determined mainly by elevation) MMPI profiles into three categories: psychosis, psychoneurosis, and conduct disorder. The classification of each profile was done rapidly—within five or ten seconds—and on an impressionistic basis. The results of this study revealed that approximately two thirds of the protocols could be accurately classified into their appropriate categories. Implicit in each of these subjective judgments, however, were certain definite but informal rules. For example, a profile typically would be classified as psychotic if it had an upward slope to

the right of the curve ($Sc > Pt$, Pa or Ma elevated), a definite peak on D, a high F, and so forth.

A more formal set of rules for classifying MMPI profiles recently has been proposed by Meehl & Dahlstrom (1960). Based on the findings of previous as well as their own research, the authors have developed and cross-validated what essentially is a "cookbook" for classifying profiles into "Psychotic", "Neurotic", or "Indeterminate" categories. Provided an MMPI meets certain validity criteria ($L < 70$, $F < 80$, and $"?" < 60$), it can be placed into one of the three classifications. The 16 proposed rules take into account the elevation as well as the configuration of scores, and each rule is applied consecutively to the K-corrected MMPI profile until a classification is reached; once a profile has been categorized by a rule, no further rules are applied. While it is not clear as to the meaning of the "Indeterminate" classification, the "Neurotic" and "Psychotic" categories presumably refer to relatively well defined diagnostic groups.

In a recent critical review of various definitions of mental illness, Scott (1958) has suggested that an important area for research lies in the investigation of the extent to which the various diagnostic measures agree and disagree in the assessment of psychopathology. Following this line, the present study was an attempt to compare the developmental scoring of the Rorschach with the Meehl-Dahlstrom rules for the MMPI as measures of severity of psychological disturbance.

METHOD

Rorschachs and MMPIs were obtained from 50 male neuropsychiatric patients, whose ages ranged from 19 to 57, with a mean age of 32.5. The criteria for selection were: (a) The MMPI and Rorschach had to have been administered within one week of each other. (b) The scores on the MMPI validity scales had to be below

certain cut-off points ($L \leq 6$, $F \leq 80$, $K \leq 70$).

The Rorschach protocols were scored according to developmental level, and Becker's (1956) rankings of the scores were used to achieve an overall developmental score for each subject. Thus, a response of the highest level of articulation and integration (W_{+}) received a rank of 6, while a response representing one of the lowest developmental levels (W_{-}) received a rank of 1. Intermediary ranks were assigned to responses falling into scoring categories between these two polar points. Each MMPI profile was classified into either the "Psychotic", "Neurotic", or "Indeter-

minate" category according to the Meehl-Dahlstrom rules (Meehl & Dahlstrom, 1960).

RESULTS

The mean K-corrected MMPI profile for all 50 Ss is presented in Figure 1. A classification of this composite profile by means of the Meehl-Dahlstrom rules places it into the "Indeterminate" category. Figure 2 presents the mean K-corrected MMPI profile for the "Neurotic" ($N = 17$), "Psychotic" ($N = 14$), and "Indeterminate" ($N = 19$) categories. As one might expect, the mean "Neurotic" profile as categorized by the rules corresponds to what typically is con-

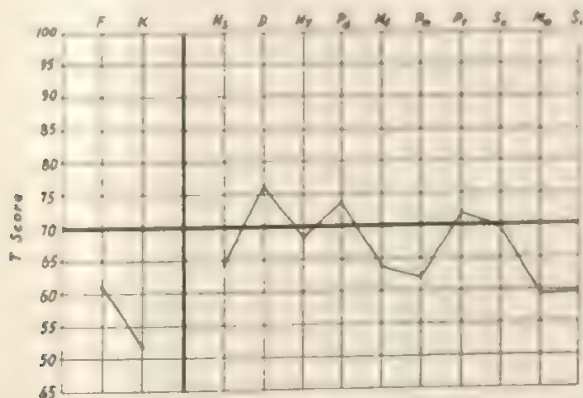


FIG. 1 Mean K-Corrected MMPI Profile for All 50 Ss.

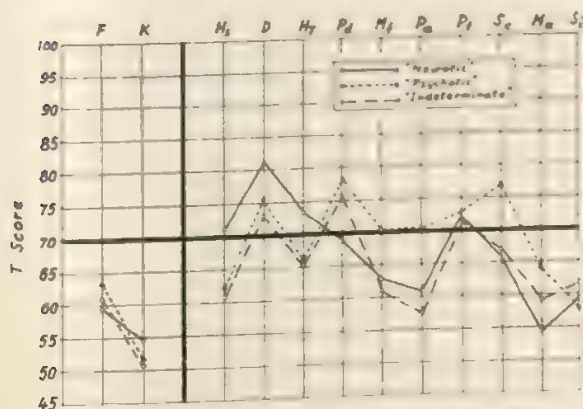


FIG. 2 Mean K-Corrected MMPI Profile for Each of the Three Meehl-Dahlstrom Rule Classifications.

sidered neurotic by impressionistic judgment. That is, there is an elevation of the neurotic triad (Hs, D, and Hy), with a secondary peak at Pt. The mean "Psychotic" profile depicts the typical upward slope to the right of the curve (Pa, Pt, Sc) with elevations on D and Pd. The profiles classified as "Indeterminate" show characteristics of both the neurotic and psychotic curves. The "Indeterminate" profile in Figure 2 (as well as the composite profile in Figure 1) resembles the "Psychotic" curve with respect to the neurotic scales (to the left of the curve) and corresponds to the "Neurotic" record in regard to the psychotic scales (to the right of the curve).

Ss with "Neurotic" and "Psychotic" profile types were compared with respect to their developmental level on the Rorschachs. The median developmental level was determined for both the "Neurotic" and "Psychotic" groups combined, and the frequency of scores falling above and below this median was obtained for each of the two groups. Table I presents the results of the median test (Siegel, 1956). The obtained chi square failed to reject the null hypothesis, in-

to measure a similar variable, viz., severity of psychological disturbance. That is, no significant relationship was found between the Rorschach's estimate of developmental level of personality organization and the MMPI's classification of neurotics and psychotics. It should be noted that this study is not intended to be a validity study for either of these two techniques, but rather represents an attempt to investigate the relationship between these measures of severity of pathology.

A question might be raised as to the nature of the profiles classified as "Indeterminate", as profiles categorized as such seem to have characteristics of both the neurotic and psychotic records. Whether protocols in this category should have been more properly classified as "Neurotic" or "Psychotic" (or perhaps even "normal") is difficult to say. As the "Indeterminate" profile appears to resemble in part both the typical neurotic and psychotic profiles, it might be an interesting investigation in itself to determine the composition of Ss that fall into this category. In the present study, however, even with the utilization of only the neurotic and psychotic profiles, no difference in developmental level was found.

A possible explanation for what appears to be a surprising inconsistency is that each of these two instruments may be measuring somewhat different aspects of functioning. Whereas an individual's perceptions of relatively unstructured ink blots manifest the quality of his organizational ability and his level of cognitive functioning, his responses to questions on a paper and pencil test such as the MMPI reflect more of his observable behavior and symptomatology. While it undoubtedly is true that an individual's level of cognitive functioning and his overt behavior are interrelated, predictions and statements made independantly on the basis of each of these at a particular point in time may be very different. Valid

TABLE I—Median Test for Developmental Level of "Neurotic" and "Psychotic" Subjects

Subjects	Above median	Below median	Total
"Neurotic"	8	9	17
"Psychotic"	6	8	14
Total	14	17	
Chi square = 0.02 df = 1 N.S.			

dicating there was no statistically significant difference in Rorschach developmental level between Ss classified by the MMPI Meehl-Dahlstrom rules as "Psychotic" and those categorized as "Neurotic".

DISCUSSION

The results of the present study failed to demonstrate a relationship between two instruments purporting

statements resulting from the assessment of a given aspect of functioning (e.g. cognition) may not necessarily correspond with valid statements concerning other aspects of functioning (e.g. overt behavior).

Before one can test the validity of either the developmental scoring of the Rorschach or the Meehl-Dahlstrom rules in the diagnosis of severity of pathology, it first may be necessary to specify the particular level of functioning in question. While one might expect the developmental scoring of the Rorschach to be more efficient than the MMPI in detecting pathological functioning before it manifests itself in overt behavior, this may not necessarily be the case; there is some indication that the Meehl-Dahlstrom rules frequently detect pathology before it becomes obvious clinically. Based on data collected from several clinical sources, Meehl² states that after a follow-up period of five years, the Meehl-Dahlstrom rules were found to have been more accurate in the original diagnosis of schizophrenia than were the clinical judgments. It might be noted in this regard that Peterson (1954), using several psychotic signs on the MMPI (e.g. Sc > Pt, Pa or Ma > 70, was able to increase the accuracy of detecting cases of incipient schizophrenia before the diagnosis could be made clinically.

The discriminating efficiency of a clinical instrument becomes especially important in cases that are difficult to diagnose. To the extent that a diagnostic instrument is valuable, it is not very valuable if it can only differentiate groups that are readily distinguishable *without* the instrument. With respect to the Meehl-Dahlstrom rules and the Rorschach developmental level, then, it might be fruitful to study longitudinally the relative efficiency of these two techniques in detecting pathology before it becomes

more evident. The implication from Meehl's findings is that the Meehl-Dahlstrom rules are more efficient than the Rorschach developmental level. Longitudinal research is needed in order to compare the discriminative power of the MMPI with that of the Rorschach developmental level, as well as with the efficiency of other diagnostic instruments. In fact, a systematic ordering of the various clinical instruments according to their discriminative ability over a period of time might offer a new approach in investigating the sequence of progression of pathological functioning.

SUMMARY

The Rorschach and MMPIs of 50 male neuropsychiatric patients were used to assess severity of pathology. The Rorschach protocols were scored according to a developmental scoring system based on Werner's genetic theory of cognition, and the MMPI profiles were classified into "Neurotic" and "Psychotic" categories by means of the Meehl-Dahlstrom rules. The results of this study failed to demonstrate any significant relationship between the Rorschach developmental level and the Meehl-Dahlstrom rules. The possibility that different levels of functioning are being measured by these two instruments was offered as an explanation for these findings, and future research along these lines was suggested.

REFERENCES

- Becker, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, 53, 229-236.
- Fine, H. J., & Zimet, C. N. Process-reactive schizophrenia and genetic levels of perception. *J. abnorm. soc. Psychol.*, 1959, 59, 83-86.
- Friedman, H. Perceptual regression in schizophrenia: an hypothesis suggested by use of the Rorschach Test. *J. proj. Tech.*, 1953, 17, 171-185.
- Gough, H. G. Diagnostic patterns on the MMPI. *J. clin. Psychol.*, 1946, 2, 23-37.
- Hemmendinger, L. Perceptual organization and development as reflected in the struc-

² P. E. Meehl, Personal communication. November 3, 1958.

- ture of the Rorschach Test responses. *J. proj. Tech.*, 1953, 17, 162-170.
- Meehl, P. E. Profile analysis of the MMPI in differential diagnosis. *J. appl. Psychol.*, 1946, 30, 517-524.
- Meehl, P. E., & Dahlstrom, W. G. Objective configural rules for discriminating psychotic from neurotic MMPI profiles. *J. consult. Psychol.*, 1960, 24, 375-387.
- Ruesch, J., & Bowman, K. Prolonged post-traumatic syndromes following head injury. *Amer. J. Psychiat.*, 1945, 102, 145-163.
- Schmidt, H. O. Test profiles as a diagnostic aid: the Minnesota Multiphasic Inventory. *J. appl. Psychol.*, 1945, 29, 115-131.
- Scott, W. A. Research definitions of mental health and mental illness. *Psychol. Bull.*, 1958, 55, 29-45.
- Siegel, E. L. Genetic parallels of perceptual structurization in paranoid schizophrenia: an analysis by means of the Rorschach technique. *J. Proj. Tech.*, 1953, 17, 151-161.
- Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- Sullivan, P. L., & Welsh, G. S. A technique for objective configural analysis of MMPI profiles. *J. consult. Psychol.*, 1952, 16, 383-388.
- Werner, H. *Comparative psychology of mental development*. (Rev. Ed.). Chicago: Follett, 1948.
- Werner, H. The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development*. Minneapolis: Univer. of Minn. Press, 1957. Pp. 125-148.

Received October 16, 1961

The Cognitive Functioning of the Creative Person: A Development Analysis¹

CHARLES HERSHEY
Cambridge Guidance Center

Recent years have shown an increased interest in the study of creativity. Questions have been asked concerning the personality of the creative person, his motivation, his distinguishing ego processes, his mental activities during the process of creation, and his general cognitive characteristics. The present study falls within the last-mentioned category, and is concerned with the pattern of cognitive functioning which characterizes the creative individual.

The conceptual framework of the investigation is Werner's (1957a) comparative developmental theory. The fundamental law of this theory states that development is characterized by increasing differentiation and hierarchic integration. Hence, when any aspect of functioning is described according to its degree of differentiation and integration, it may be given a relative developmental position. The theory, then provides for a description of the formal properties of cognition, and permits a comparison, among groups, of cognitive processes which are categorized developmentally.

Werner (1957b) makes the assumption that each individual, rather than operating at a fixed level of functioning, continually varies in the genetic level of his performance. However, he suggests that the ability to utilize processes at varying developmental levels is of particular significance in creativity. He states:

Now, creativity, in its most general meaning, is an essential feature of emergent evolution, and this, in turn, implies progression

through reorganization. Since we assume that such progress through reorganization cannot be achieved without "starting anew", that is, without regression, it follows that a person's capacity for creativity presupposes mobility in terms of regression and progression. (Werner, 1957b, p. 145)

Creating, in this light, is seen as a continual repetition of a particular biphasic process. In the first phase, there is a breaking down or dedifferentiation of previously existing patterns. Following the regressive phase, there is a phase of progression in which the reorganization of elements into previously non-existing patterns take place.

Such developmental shifting implies that the individual has available in his repertoire responses that are representative of both genetically mature and genetically primitive levels. Although all individuals are assumed to vary to some extent in the level of their functioning, the creative person has been distinguished for an activity that is said to have as its basis considerable developmental shifting. The implication is that creative and non-creative persons differ in the degree to which they have available varied genetic processes. The hypothesis advanced here may be stated as follows: In any situation in which responses can be developmentally ordered, creators will show a more ready availability of both relatively mature and relatively primitive responses than noncreators.

A second concern of this study is with the character of the primitive functioning of creators. In this regard, it may be noted that there are pathological groups, particularly schizophrenics, who are also expected to evidence primitive processes to an unusual degree (Friedman, 1952; Sie-

¹This study is based on a doctoral dissertation done at Clark University. The author wishes to express his appreciation to Leslie Phillips, Bernard Kaplan, and Gordon Gwinn for their many contributions.

gel, 1953; Werner, 1957a). While it is expected that mature responses are more readily available to creators, the question is raised as to whether the primitive functioning of creators differs from that of schizophrenics.

METHOD

Three groups of subjects, each composed of 20 male adults, were studied: creators, noncreative normals², and schizophrenics. The relative primitivity and maturity of their functioning was evaluated on the basis of a developmental scoring of Rorschach protocols. For purposes of this study, an empirical criterion of creativity was used. Hence, by "creative individual" is meant a person who has achieved prominence as a creator in one of the major cultural domains.

Subjects

The *creative group* consisted of a sample of eminent artists, whose cultural contributions have been generally acknowledged.³ The *normal group* was composed of individuals who were in no way distinguished for creativity in any of the culturally ascribed forms, and for whom there was no evidence of psychosis. Regarding occupational status, there were nine industrial workers, five firemen, two insurance salesmen, two entrepreneurs, one industrial executive, and one physician. The *schizophrenic group* consisted of hospitalized patients who carried a diagnosis of Schizophrenia or Dementia Praecox. Fourteen of these subjects were of the Paranoid type, while the remaining six were distributed among the other sub-classifications.

The subject groups were selected so as to be relatively homogeneous with

respect to age, intellectual level, and Rorschach response productivity. There were no statistically reliable differences among the groups on these factors ($P = > .10$ in all instances).

Assessment of Cognitive Functioning

The Rorschach test (individually administered to each subject) was the instrument used to assess cognitive functioning. Responses to this test can be classified developmentally by means of the Genetic Scoring System (Phillips, Kaden, & Waldman, 1959). In this system, responses are categorized as relatively mature or primitive on the basis of formal properties described in Werner's developmental theory. Hence, responses indicating differentiation, articulation and integration are considered mature, while those characterized by diffuseness and syncretism are considered primitive.

In this investigation, six response categories were studied: Movement Response, Integrative Response, Form Dominant Response, Form Subordinate Response, Physiognomic Response, and Primitive Thought Response.⁴ All the response categories except the Physiognomic Response (which is presented here for the first time⁵) have been validated as to their developmental classifications by ontogenetic studies (Ames, et. al., 1952; Ford, 1946; Hemmendinger, 1953; Ledwith, 1952).

The Movement Response (M): The Movement Response contains the traditional human movement scores. It involves the ascribing of action to the static forms of the inkblots. Such motion perception implies that organismic behavior is occurring cognitively rather than motorically (Wer-

⁴A detailed description of the scoring criteria for these responses, except for the Physiognomic Response, may be found in Phillips, Kaden, and Waldman (1959). Complete criteria for scoring the Physiognomic Response may be found in the author's dissertation (Hersch, 1957); or a copy may be obtained from the author.

⁵A response called the "Physiognomic Dd" by Siegel (1953) is different from the Physiognomic Response presented here.

²From this point on, the terms "normals" and "noncreative normals" will be used interchangeably.

³We are indebted to Anne Roe for making available from her research files the data for the creative group. For an indication of the professional standing of these artists, see Roe (1946).

ner, 1945), with the delay of immediate motoric responsivity indicating self-environment differentiation. Hence, this response is considered mature.

The Integrative Response (I): The Integrative Response involves the appropriate organization of two or more adequately perceived (F+) blot elements into a unified response. The relative articulation, differentiation, and hierarchic integration in such responses are indicative of developmental maturity.

The Form Dominant Response (FD): This category includes those responses in which form is dominant over color or shading (including achromatic color, vista and texture). These responses tend to be relatively articulated and differentiated, and indicative therefore of genetic maturity.

The Form Subordinate Response (FS): This category includes those responses where color or shading is primary, and form is secondary or not used at all. These responses tend to be relatively diffuse and undifferentiated, and therefore developmentally primitive.

The Physiognomic Response (PR): This category includes those responses in which the blots themselves, and not something seen on the blots, are considered as having affective or symbolic qualities, dynamic properties, the attributes of living things and the like. The usual interpretation of the blot as "looking like" a certain object or thing is minimal, and the blot itself is described as being "gay", "vicious", "troubled", as having "undulating movement", and so on. The responses in this category reflect the genetically primordial activity termed "physiognomic perception" (Werner, 1956; Werner, 1957a). They are determined by a fusion of the motoric or affective attitudes of the subject with the sensory stimulation provided by the blot. Such responses reflect a syncretic relationship between the self and the external world, and are therefore considered primitive.

The Primitive Thought Response (PT): This category includes the usual contaminations, tabulized combinations, and confabulations. In the first two there is a syncretic fusion of elements into an arbitrary final percept. The confabulations reflect the diffuse association of part and whole ("pars pro toto" functioning). The Primitive Thought Category, then, is considered primitive.

In summary, the Movement, Integrative and Form Dominant Responses are considered mature; the Form Subordinate, Physiognomic, and Primitive Thought Responses are considered primitive.

The writer scored all of the Rorschach protocols, but to insure precision in scoring any response that raised a question was submitted to at least one, and sometimes to two, judges for checking. In all such cases, unanimity of scoring was reached.

HYPOTHESES AND RESULTS

In the earlier discussion, certain general hypotheses were stated, and these will now be translated into specific propositions. The frequency of responses in the various Rorschach categories was used as the index of the availability of mature and primitive cognitive processes. Comparisons between groups with respect to frequency in a given response category was carried out by means of the Mann-Whitney two-sample test (Mann and Whitney, 1947). In those instances where directional predictions were made, one-tailed significance tests were used. Inasmuch as this study is in a relatively unexplored domain, a significance level of .10 was selected for all tests of hypotheses. Medians and comparisons for all groups may be found in Table I.

The major hypothesis was that creative individuals have a greater availability of both mature and primitive operations than non-crea-

⁸The writer is indebted to Marvin Waldman and Leslie Phillips for acting as judges.

Table I. Medians and Statistical Comparisons Among Groups on the Rorschach Categories

	Mature Categories			Primitive Categories		
	M	I	FD	FS	PR	PT
Artists	4	5	6.5	3	2	1
Normals	2	6	5	2.5	0	0
Schizophrenics	1	2	2	3.5	0	1
	Comparisons ^a					
	M	I	FD	FS	PR	PT
	Direction P	Direction P	Direction P	Direction P	Direction P	Direction P
Artists—						
Normals	A>N .05	— —	A>N .10	— —	A>N .01	A>N .10
Artists—						
Schizophrenics	A>S .01	A>S .01	A>s .01	— —	A>S .01	— —
Normals—						
Schizophrenics	N>S .05	N>S .01	N>S .05	— —	— —	S> .05

^a All one-tailed tests except Artist-Schizophrenic comparisons on the primitive categories.

tive normals. This may be translated into two empirical hypotheses.

Empirical Hypothesis I. The artists will give more Movement, Integrative, and Form Dominant Responses than the normals. This was corroborated for the Movement Response and the Form Dominant Response.

Empirical Hypothesis II. The artists will give more Form Subordinate, Physiognomic, and Primitive Thought Responses than the normals. This hypothesis was borne out for the Physiognomic and Primitive Thought Responses.

In short, four of the six predictions in comparing artists and noncreative normals were supported by the findings. Predictions concerning the Integrative and Form Subordinate Responses were not supported.

In comparing creators with schizophrenics, the expectation was that mature responses are more readily available to creators.

Empirical Hypothesis III: The artists will give more Movement, Integrative and Form Dominant Responses than the schizophrenics. This hypothesis was supported by the findings for all three of the mature categories.

No hypotheses were stated in regard to the primitivity of creators as compared with schizophrenics. This was evaluated by means of com-

parisons between the artists and schizophrenics on the primitive Rorschach scores. It was found that the artists gave reliably more Physiognomic Responses than the schizophrenics; neither the Form Subordinate Response nor the Primitive Thought Response statistically discriminated between the two groups.

Although not directly involved in our formulations, it was of interest to compare the normals and schizophrenics on the Rorschach categories. One would expect the normals to reveal a greater availability of mature processes, and the schizophrenics to reveal a greater availability of primitive processes. As seen in Table I, this was borne out for all of the mature Rorschach scores. In regard to the primitive scores, the schizophrenics gave more Primitive Thought Responses, but no reliable differences were found for the Physiognomic Response or the Form Subordinate Response.⁷ The results which support the expectations corroborate previous findings and conceptualizations.

DISCUSSION

The major hypothesis of this study was that the cognitive functioning of

⁷It will be noted that the Form Subordinate Response did not discriminate among the three groups; it will not be considered further.

the creative individual, as compared with the noncreative normal, is characterized by the greater availability of both relatively mature and relatively primitive processes. The findings tend to support this hypothesis. In regard to mature operations, creators have available more responses reflecting differentiation and articulation; in regard to primitive operations, creators have available more responses indicative of syncretism and diffuseness. The creator, therefore, more so than the noncreative normal, demonstrates the flexible, mobile utilization of both genetically advanced and genetically early functioning. Hence, the findings give credence to Werner's theoretical suggestion that "a person's capacity for creativity presupposes mobility in terms of regression and progression."

To turn to the comparison of the artists and schizophrenics, a striking finding is that the artists responded more frequently with the Physiognomic category. These responses are in fact rare among schizophrenics, with only six of them being given by the total group in this study. In addition, the author's experience with the Rorschachs of children indicates that these responses are quite rare for them as well. The Physiognomic Response, then, appears to reflect primitivity of a special character.

It seems plausible to attribute the special character of this primitivity to a certain mode of functioning which is utilized by the creator, but which is not readily available to the schizophrenic. It is known that the effects of a syncretic self-environment relationship are often noted in schizophrenia, frequently resulting in physiognomic perceptual phenomena (Werner, 1957a). Yet these phenomena are not frequent in the schizophrenic's Rorschach performance. There would seem, then, to be an important difference between the functioning of a schizophrenic who sees an open door as a sinister, devouring mouth from which he flees

in panic, and the functioning of a creative person who, in responding to the Rorschach, states:

Most interesting of all besides the physical description of certain silhouettes I get a certain feeling I felt in childhood when I entered a little country town, children would play in an abandoned house, the feeling of horror or uncertainty, a haunted quality. You can make out the forms in a dimly lit room and the child imagines inner terrors. The feeling is quite terrible. That's the thing that comes to my mind. Unknown form. Like looking into the unknown. A supernatural feeling. A child likes to look at those things. Today I don't bother with them. Something terrifying about the symmetry. This damn thing being again over here. And the thing quivers like a spectrum . . . the forms making a boundary so the quivering seems to be in a contained space, but it is throughout the blot.

In the case of the schizophrenic, the fusion of subject and object gives rise to physiognomic perceptual phenomena that are then given the status of an objective reality which is reacted to accordingly. In the case of the creator, the fusion also gives rise to physiognomic perceptual phenomena, but the creator is then able to reflect upon the experience, objectify it, and distinguish between what is within himself and what is valid for the external stimulus. Such a second step implies self-environment differentiation.

In short, the Physiognomic Response is considered to involve a biphasic process, first a regression and then a progression. And it will be recalled that a similar biphasic process was presumed to underlie creative activity. It is perhaps not surprising, then, that creators offered the Physiognomic Response more frequently than either noncreative normals or schizophrenics, nor that this category failed to discriminate between the two latter groups. In actuality, it is this response, more than any other, which distinguishes the Rorschach protocols of creators from those of the other two groups. In this study, only 6 such responses were found in

the total normal group and in the total schizophrenic group. 49 such responses were given by the artists.

Therefore, in terms of the second major concern of this study, it does appear that there are differences in primitivity between the creator and the schizophrenic. For one thing, the creator is able to subordinate and utilize his primitive operations in a manner that is rare in schizophrenia. Secondly, the primitive functioning of the creator is more circumscribed, with mature operations readily available (M.I.F.D.), while the schizophrenic is relatively limited to genetically early functioning. It is as if in the one case the primitivization dominates the overall functioning, while in the other there is a ready shift to a more mature performance. To conceptualize more broadly, we might say that for the creator primitive operations may be exploited toward the production of the highest cultural achievements, while for the schizophrenic primitive operations serve the pathological purposes of symptom formation (delusions, hallucinations, etc.).

Further consideration may be given to the question of what enables the creator to indulge in primitive functioning to the extent that he does. Some of the findings in the present study may throw light on this question.

In traditional Rorschach interpretation, both the Form Dominant Response and the Movement Response are said to indicate some aspect of organismic control (Beck, 1945; Klopfer & Kelley, 1946; Rorschach, 1942). Although it need not be elaborated here, it may be noted that these traditional interpretations can be readily coordinated with the genetic interpretation of the responses offered earlier. The findings for the M and FD responses, therefore, indicate that controlling functions tend to be more readily available to creators than to noncreative normals. It is possible that the difference in degree makes

for a difference in quality, with the creator having optimal or flexible controls, and the noncreative normal having rigid or self-limiting controls. In any event, regarding the creator's more frequent expression of primitive operations, the following formulation is offered. The creator, by virtue of his available controlling operations, is able to shift frequently into relatively primitive modes. This shift is temporary and, as has been noted, the creator not only performs readily at a genetically advanced level again but objectifies and utilizes the primitive process by imposing upon it developmental operations of a mature nature.⁸ The noncreative normal, on the other hand, lacking the degree or kind of control found in creators, is restricted in his ability to function primitively. The schizophrenic, it may be added, even further lacking in controls, is overwhelmed by his primitive functioning.

Noting such differences in primitive functioning makes one take pause and consider the question of the "value" of primitivity. It has been reiterated (Werner & Kaplan, 1956) that for developmental theory the concept of primitivity has no evaluative connotation. Nevertheless, value judgments often do enter into the consideration of genetically ordered functioning, with primitivity labeled as "bad" and maturity extolled as "good". The present study indicates that such a generalization has to be reconsidered. In other words, if value judgments are to be made at all, the potential positive value of primitive functioning should be considered as well as the negative.

A number of other writers share this point of view. Werner and Kaplan (1956), for example, point to the importance of primitive processes for such culturally highly regarded activ-

⁸Analogous formulations are proposed by Kris (1944; 1952) whose concepts of "regression in the service of the ego" or "controlled regression" in themselves imply a duality of genetic functioning.

ities as esthetic enjoyment and adequate interpersonal relationships. They found that a primitive orientation in the structuring of verbal concepts involves the utilization of genetically earlier cognitive functions than a primitive orientation. Maslow (1950) states that "self actualizing" people are capable of greater de-differentiation of ego boundaries than others would consider possible. Shor (1954) discusses the importance of "non-neurotic regression" for the maximal fulfillment in sexual love. Kris (1944; 1952) has placed major emphasis on the role of primitive functioning in artistic creativity, and his concepts have been employed by Rosen (1953), who discusses the role of primitive modes of perceiving space and number in the attainment of mathematical insights. These examples do not exhaust the list, but are sufficient to indicate some agreement with the position taken here: before any instance of primitive functioning can be regarded as "good" or "bad", some consideration has to be given to the context in which it occurs and the products to which it gives rise.

SUMMARY

This study, based on Werner's developmental theory, compared the genetically ordered cognitive functioning of creative people (represented here by a group of artists), non-creative normals, and schizophrenics. The cognitive instrument employed was the Rorschach test. In support of the major hypothesis, it was found that the cognitive functioning of the creator, as compared with the non-creative normal, tends to be characterized by the greater availability of both relatively mature and relatively primitive processes. In comparing creators and schizophrenics, it was found that creators have a more ready availability of mature processes; also, these two groups tend to differ in the nature of their primitive functioning. Differences among the three

groups in the expression of primitive functions was found in the relative ability of control functions. Consideration was given to the question of the "value" of primitivity.

REFERENCES

- Abel, Louis E. *Emotion and Memory*. New York: Wiley, 1949.
- Beck, S. J. *Rorschach's test, Vol. 2, A variety of personality pictures*. New York: Grune & Stratton, 1942.
- Freud, Sigmund. *The application of the Rorschach test to young children*. *Monographs of the Psychoanalytic Society*, 1924.
- Friedman, H. Perceptual regression in schizophrenia. An hypothesis suggested by the work of the Rorschach test. *J. Gen. Psychol.*, 1952, 47, 63-80.
- Hemmendinger, L. Perceptual regression and regression in content in the content of the Rorschach test responses. *J. Gen. Psychol.*, 1952, 47, 81-90.
- Hersch, C. The significance of the structure of the Rorschach test. A preliminary report on the results of the Rorschach test. Unpublished doctoral dissertation, Clark Univ., 1957.
- Klepper, B. & Kelley, D. M. *The Rorschach test*. New York: Holt, 1946.
- Kris, E. Art and regression. *Trans. N.Y. Acad. Sci.*, 1944, 6, 200-202.
- Kris, E. *Psychoanalytic explorations in art*. New York: The Arts Press, 1952.
- Ledwith, Nettie H. Rorschach responses of the schizophrenic and normal. Progress report. *J. Proj. Tech.*, 1952, 16, 50-51.
- Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Stat.*, 1947, 18, 50-60.
- Maslow, A. H. Self actualizing people: A study of psychological health. *Psychological Symposium #1*, 1950, 11-34.
- Phillips, L., Kaden, S., & Waldman, M. Rorschach indices of developmental level. *J. genet. Psychol.*, 1959, 94, 267-285.
- Raeff, I. G. The effect of formal and poetic orientations on the meaning structure of words. Unpublished doctoral dissertation, Clark Univ., 1956.
- Roe, Anne. Alcohol and creative work. *Quart. J. Stud. Alcohol*, 1946, 6, 415-467.
- Rorschach, H. *Psychodiagnostics*. New York: Grune & Stratton, 1942.
- Rosen, V. H. On mathematical illumination and the mathematical thought process: A contribution to the genetic development and metapsychology of abstract thinking. *Psychoanal. Stud. Child*, 1953, 8, 127-154.

- Shor, J. Female sexuality: Aspects and prospects. *Psychoanal.*, 1954, 2, 47-76.
- Siegel, E. L. Genetic parallels of perceptual structurization in paranoid schizophrenia: An analysis by means of the Rorschach technique. *J. proj. Tech.*, 1953, 17, 151-161.
- Werner, H. Motion and motion perception: A study on vicarious functioning. *J. Psychol.*, 1945, 19, 317-327.
- Werner, H. On physiognomic perception. Unpublished manuscript, Clark Univ., 1956.
- Werner, H. *Comparative psychology of mental development*. (Rev. ed.). New York: Int. Univ. Press, 1957. (a)
- Werner, H. The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development: An issue in the study of human behavior*. Minneapolis: Univ. of Minnesota Press, 1957. (b)
- Werner, H. & Kaplan, B. The developmental approach to cognition: Its relevance to the psychological interpretation of anthropological and ethnolinguistic data. *Amer. Anthropol.*, 1956, 58, 866-880.

Received October 13, 1961.

Developmental-Diagnostic Dimensions As Seen Through Psychological Tests

ANNELIESE F. KORNER

Department of Psychiatry, Mount Zion Hospital & Medical Center, San Francisco, California.

INTRODUCTION

In our present state of knowledge about child development, valid and consistent criteria for the necessity of treatment are difficult to establish. We neither have a solid notion of what constitutes psychological health, nor do we have a good nosology of childhood disorders. We borrow heavily from adult psychopathology even though we are painfully aware that this nosology is inadequate for describing even adult patients. We share in the difficulty of being more aware of the pathogenic than of the adaptive mechanisms. This is not surprising since adaptive efforts which fail are easier to witness than successful ones, since these often go on in silence. Added to this, in the case of children, are the vicissitudes of the developmental process which frequently make it difficult to judge whether a given symptom is a transitory product of a new maturational step or the first sign of pathogenesis. Symptom diagnosis, altogether, is fraught with difficulty. A very rare and unique symptom may constitute an effort at self-healing which may succeed and subside. By contrast, common symptoms which fall within the range of statistical normality may be very crippling, depending on their specific function in a child's psychic economy.

What are we, then, to take as a workable criterion for the necessity of a therapeutic intervention? Anna Freud (1945) many years ago gave us an excellent criterion which eliminates many of the above difficulties. She pointed out that the one thing that really need concern us is a threat to, or an arrest of the developmental process itself. The task of this kind of

evaluation involves judgment regarding the intactness both of the child's libidinal and ego development. Clinically, it is a well-known fact that arrest in one segment of development often leads to a spread to other segments, and that this process can become very crippling, leaving an imprint on all future development. This typically happens in the severe childhood neuroses. It also happens when, for one reason or another, curiosity is stunted, or acquisitiveness is inhibited; that is, when for the sake of internal security new experiences are avoided. The very common learning disabilities are often an outgrowth of these ego restrictions, and in their effect probably come closest to being a direct impairment of the growth process.

We are thus called upon to judge the normality of the developmental process, and our task is one of making a developmental diagnosis. Making such a diagnosis affords us a number of advantages over other diagnostic statements: it helps us view isolated symptoms and anxieties in perspective, and leads us away from absolute standards of psychological health. It is the only type of diagnostic statement which includes genetic considerations with all the etiological hypotheses that this implies, and it relies heavily on the child's unique developmental history to spot regressions, progressions and arrests of various functions.

DEVELOPMENTAL-DIAGNOSTIC TOOLS

Our most important tool in making a developmental diagnosis is a solid grasp of child development, normal and deviant. Unfortunately, the state of our knowledge in this area is still

very incomplete; we do not as yet have a clear model of the vicissitudes of the developmental process. But we do have important guideposts: we have developmental norms by Gesell (1947), Terman (1937), and many others. We have theories of the development of thought by Werner (1947) and by Piaget (1930, 1952). And we have the psychoanalytic model of libidinal development as described most vividly by S. Freud (1953) and Erikson (1950). Intuitively, the clinician borrows from all these guideposts in understanding his case. In each case this requires a creative synthesis of these various points of view. Our task would be infinitely facilitated if we could rely on a systemized synthesis of these. Very interesting beginnings in this direction have been made by Wolff (1960) and Anthony (1956, 1957). We sorely need an extension of these. In making a developmental diagnosis we want to understand regressions, fixations and progressions not only in libidinal terms but also in ego terms, such as through Piaget's stages of thought and Werner's levels of differentiation. A systematic linkage between affective and cognitive genesis has not only very important clinical implications, but is also of tremendous theoretical interest. Clinically, we might ask ourselves whether an unresolved maturational crisis expresses itself in characteristically arrested modes of thought. Possibly one might date the onset of deviant development from an analysis of the thinking process and the extent, depth and nature of its impairment. Theoretically, Piaget's epistemology may be helpful in establishing a genetic hierarchy of the defenses, a problem which psychoanalysis has great difficulty solving.

The developmental psychologies of Piaget, Werner and psychoanalysis thus can supply useful yardsticks for the evaluation of the level of maturity of a child's cognitive and affective development. This is true also of psychological tests.

TESTING AS AN AID IN DEVELOPMENTAL DIAGNOSIS

Psychological tests can contribute a great deal in gauging the maturational level of a child's ego, libidinal and superego development. The remainder of this paper aims to illustrate this point.

1. Evaluation of Ego Characteristics. Psychological tests are ideally suited for revealing certain ego characteristics. Ordinary cognitive tests can give us a great deal more than an intelligence rating. We need not dwell here on the gross diagnostic clues suggestive of psychosis or intracranial pathology. The latter becomes quite clear through signs of concrete and/or perseverative thinking, an inability to shift and to revise errors, perplexity, primitivization and visual and motor disturbances. Psychosis is usually spotted readily through evidence of bizarre ideation, irrelevancies and other forms of thought disorder. More subtly, intelligence tests can be used to pinpoint developmental lags and focal precocities of function. A great mass of norms backs our clinical impression of how even or uneven a child's cognitive development is. From the variability of performance one can gather how reliably or unreliably certain ego functions work. Variability in function is a good measure of cathectic changes: the more concern there is with intrapsychic problems at any given moment the less energy is available for tackling intellectual tasks realistically.

Case Illustration. The following is an example of the kinds of diagnostic clues one can obtain purely from a cognitive test. In this instance, the intelligence test revealed not only many of this boy's ego characteristics, but also arrests in his libidinal development. It furnished many specific clues for a developmental diagnosis which could be further explored through other means.

Bob was referred to me with a

minimum of clinical information. He was a twin, 12 years old, obese, chatty and old-maidish in manner. By talking incessantly from the start, he kept on the initiative and could control to some extent what was going to happen to him. One could sense from this and many other clues that he feared passivity. Judging from the parents' professions they were of superior endowment. There was no mention that Bob had any scholastic difficulties. Thus, presumably, we were dealing with a bright boy.

Bob was nevertheless given an I.Q. test because his Rorschach and other observations suggested an advanced stage of an obsessional state which only thinly covered all sorts of paranoid suspicions. To be explicit: he gave nothing but vague and noncommittal responses, but even these were too much of a commitment for him. He immediately doubted what he said and took it back. An inquiry into the nature of his responses was experienced as an ordeal by him; when requested to explain his answers he felt that he was irrevocably pinned down. He behaved as if he were going to be brought to trial and crucified for what he thought might turn out to be "lies." As he felt pinned down he kept wondering why my office did not have any windows. From all this evidence I suspected not only a near-psychotic disturbance, but also a severe paralysis of his intellectual functions. He was given the Revised Stanford Binet, Form L, and my suspicions were confirmed: he earned an I.Q. of 88 which is considered dull normal. The range and succession of his passes and failures suggested not only that this was a low estimate of his abilities, but also testified to the unreliability of many of his ego functions. His performance was extremely unpredictable; in certain instances he passed more items at high levels than at

low levels. He was 12 and yet he failed some items expected of 8-year-old children, and passed others expected of 14-year-olds. His successes at high levels were, in themselves, an indication that basically Bob was not dull.

Bob's judgment was considerably impaired. The very fact that he was not insulted when asked to tackle problems which 7-year-old children readily solve suggested this. At the same time, he inappropriately remarked that items at the Average Adult level were easy, showing no awareness of his failures. Furthermore, anxiety clouded his judgment. He magically undid and thereby denied and misjudged situations which implied any bodily harm to anyone. In some of these situations he tried to create distance for himself by saying, "Are these stories real?" In responding to absurd stories he contended that a *well-known* railroad could not possibly have an accident; he insisted a hanging was just a threat even though the story suggested otherwise. One could not possibly find the skull of a 10-year-old. It was inconceivable to find a man bound up in his chair. These situations, which ordinarily create interest in the Binet, were so close to threats he felt acutely that he had to turn them around for comfort. One could sense undoing in his self-critical comments also; by saying that he failed a problem he hoped that perhaps it was not so.

Both his work approach and the content of his answers suggested marked oral fixations. His ways of dealing with stimuli reflected oral immaturity: he was quite incapable of delay; he usually "bit off" only parts of a problem, being unable to take in its totality. He became acquainted with new materials by smelling at them. He ate his own nosepickings, especially as lunch time neared. He responded with

oral fantasies to an absurd story in which it was suggested that since the last car of a train usually gets damaged most, it be taken off before the train starts. He said, "The caboose might be the most important car of the train. They might cook in there. People might get hungry. At least that's my point of view." And judging from his obesity, it was. When asked what happened at a neighbor's house who was visited first by a doctor, then a lawyer and then a priest, he answered, "They all started talking at once." This was a clear projection of his own need in this direction.

In looking at pictures Bob demonstrated an impaired sense of space and perspective. In one instance he could not differentiate between bushes and clouds. He interpreted a man walking toward the horizon as "walking to the edge of the sunset and then dropping off." This suggested not only distortion in form but also a sense of doom in content. His motor coordination was so poor that he barely was able to copy a diamond, a task which is expected of 7-year-old children. All these indices of disturbed motility and perception not only attested to the depth of his disturbance, but also raised the question of an organic deficit. Looking at the evidence from another vantage point, however, it was felt that in his case what looked organic may have had to do with feeling incomplete and awkward without his twin. Bob was left-handed, his twin was right-handed. He described himself and his twin as "three-fourth the same," as if he were speaking of one unit. When asked to remember and reproduce a design he drew a double of each line. On face value this looked like a perseverative organic-like production, but in his case in all likelihood reflected a need for his twin. This hunch was corroborated by many other observations.

He was quite unable to stand alone on any issue. He never was able to express any opinion without frantically having to find out what I thought on the subject. He expressed acute anxiety lest he knew my position. This covered all subjects; God, politics, food, my family, etc. He behaved as if some catastrophe would befall him if he were not told. When he was given closure on any of these issues he sighed his relief, saying, "Now I feel good." (Complete?) In his need not to stand alone he seemed to require physical coverage by another person. The most unmistakable example of this need to merge occurred when he refused to walk side by side to my office. Had he just walked behind me so as to avoid walking in front of me, I might have interpreted his behavior as a paranoid maneuver. What he did was to walk behind me so that it was absolutely impossible for me to see him. Every time I turned around to talk to him he disappeared out of my field of vision. He ducked and took precisely the same steps right behind me. It was as if he were my shadow or my "Doppelgänger."

If we looked at what could be culled from the intelligence test alone a good number of informative clues emerged. On the ego side we saw the level at which he functioned and also got a glimpse of his potential. We witnessed his impaired and inappropriate judgment, his disturbed motility and distorted orientation in space. Ego defenses witnessed were undoing, denial, identification and projection in that order of frequency. These defenses served to deal with conflicts which were also reflected in this test. Denial, identification and undoing were aimed at defending against castration anxiety and super-ego pressures. Projection was seen as a displacement maneuver of

his own orality. The test also pointed to a marked cathectic lability and to drive-dominated thinking. There was even a glimpse of the nature of his object relationships. He used identification and his twinship defensively. He still needed to be merged with someone else and made very little attempt to differentiate himself.

One of the most helpful characteristics of tests is that they furnish us with samples of thought, verbatim and complete, which permit a microscopic analysis of content, form and sequence of thought. Because of these properties, it is easier through tests than through interviews to witness the most permanent and the most transient aspects of ego functioning. An analysis of the formal attributes of thought processes is apt to reveal modes of handling stimuli, cognitive characteristics and perceptual styles which are very much part of the grain of the ego. At the same time, an analysis of the sequence of thought will reveal transient fluctuations in the level of ego functioning. We are dealing here with expressions of cathectic shifts, momentary regressions and efforts at resynthesis on the part of the ego. Individual variations in the depth and the proneness for such fluctuations are of added diagnostic significance.

A minute scrutiny of how a child deals with unfamiliar stimuli has particular importance for an evaluation of ego functioning. Can he permit himself to be sufficiently passive to let them in or does he feel them as painful intrusions against which he has to defend himself? Does he become flooded and disorganized by them or does he dose them so stringently that his experiences become scotomized? Does he obliterate the stimulus altogether, not seeing it? Does he avoid passive reception to the point of having to select and substitute stimuli of his own, thus staying on the initiative? Does he combine more than one mode

to form several lines of defense? One can readily imagine a child overlooking a threatening stimulus simply as a defense against being overwhelmed or disorganized by it. One can also see that under such circumstances the important function of experiencing anxiety as a warning signal would be grossly impaired in such an individual.

The importance of the origin of ego characteristics for diagnosis, prognosis and treatment. If we talk of stimuli being retained, taken in, digested, warded off, diffused, ejected, etc. we conceive of perceiving as partaking in what Erikson calls organ modes (1950). There is very little doubt that psychosexual fixations leave their imprint on perceptual style, and, by implication, on modes of thought. In all likelihood the reverse is also true: ways of perceiving and thinking influence the course of psychosexual development. Observations of very young infants certainly show marked and individual differences in sensitivity to stimuli and activity patterns. (See Escalona and Heider (1959)) and Fries. Shortly after birth we see differences in receptivity, eagerness to incorporate, in need to ward off and to take flight from stimuli. Perhaps we witness here antecedents of later modes of dealing with stimuli. Leaving aside for the moment the very important effects of experience and of accommodation to and identification with parental modes, it seems likely that these innate ways throw a very individual cast on a child's development. In particular, they should make a difference in why objectively identical experiences can be experienced differently, why they prove traumatic to some and not to others. As Benjamin (1961) says, they can even "help determine what experiences will be experienced, and how they will be perceived." Together with basic differences in drive organization and in maturational rates these innate ego variations should leave an imprint on the course of psy-

chosexual development, predispose to certain fixations (Freud, 1946), influence the choice of preferred defenses (Hartmann, 1950) and the development of preferred cognitive styles (Gardner, *et al* 1959). We may deal here with the earliest core of both affective and cognitive development, or the antecedents of both. Needless to say it is very important to study this core at this relatively unelaborated stage of development. What is more, it will be the task of longitudinal studies from infancy onward to trace the vicissitudes of these early modes on later development.

These problems of genesis have more than a purely theoretical significance. They influence diagnosis, prognosis and possibly treatment methods. We need to learn to differentiate between the products of defensive elaborations of unresolved maturational tasks posed by the various psychosexual stages and the manifestations of basic ego characteristics. This is a very difficult task since the end products of both often look very much alike, particularly when the innate ego variations are reinforced through secondary neurotic elaborations. To be more concrete, there is a difference if a child's hyperactivity is an expression of what Fries calls his "congenital activity type" or a function of a frantic fear of passivity and of castration anxiety, or whether one is superimposed on the other. There is also a difference between the child who from birth has high stimulus barriers which limit his sensory experiences and the child who, for reason of fear of curiosity originating from experiences in the oedipal phase, has a reduced capacity to see, to hear and to take in. Theoretically, it is easy to see how the first child might be predisposed to experience the onslaught of the oedipal phase as potentially traumatic simply because he is already handicapped in integrating stimuli of any variety and intensity. He may therefore neurotic-

problem long before the advent of neurotic conflict. Our therapeutic methods are designed to deal primarily with the products of neurotic conflict and the neurotic elaborations of ego deviations rather than with the modification of basic ego characteristics. Often treatment alters the former and what seems to be residuals are expressions of the latter. These may require methods other than psychotherapy as we know it, methods which are more pedagogical in nature and geared to the practice, the expansion and the strengthening of certain ego functions.

Like all cross-sectional observations, psychological tests are ill-equipped to differentiate between the differences in origin of a particular difficulty and to reconstruct their genetic development. (For a more extensive discussion of the scope and limitations of psychological tests see Korner (1956)). Nevertheless, one can get hunches about these differences in origin through repeated evaluations pretherapy and posttherapy, judging, at least, from my experience with child analytic cases. It has happened that I suspected an innate predisposition for a given ego characteristic from its invariability, chronicity and indiscriminate nature, only to find basic alteration due to the analysis of certain conflicts. From these changes one can infer that the particular difficulty was more invested with neurotic conflict than was originally suspected. In other cases, the opposite might be conjectured when we find no changes on second evaluation, in spite of competent and prolonged psychotherapy of what, on first inspection, looked like classically neurotic conflicts.

While test evaluations, taken singularly, rarely shed light on genetic developments, they are capable of reflecting their end products. They are particularly suitable for study of the specific ways a given child's ego operates, which in itself holds implications for the kinds of obstacles there

might be in the way of treatment. Too little has been said about how cognitive, motor and perceptual habits can implement the most stubborn resistances to treatment, irrespective of the total diagnostic picture. The work of Gardner, *et al* (1959) on cognitive styles opens entirely new vistas in this respect. While the common psychological tests yield much grosser observations than do the laboratory procedures of Gardner, *et al*, they can, nevertheless, alert us to many aspects of cognitive habits and perceptual modes.

A case in point is a type of scrutiny which has already been alluded to, namely, a systematic analysis of how a child deals with unfamiliar stimuli. For example, whether a child takes in, ejects, wards off, or substitutes his own stimuli has tremendous repercussions for his capacity to learn. An analysis of this type is probably one of the most meaningful steps that can be taken in the evaluation of children with learning problems. It can give us valuable clues about a child's self-dosing devices and along with these some glimpse of his preferred discharge patterns. Such an analysis may also have direct implications for what may be a desirable treatment technique: it may suggest how much structure a child may need or how much freedom he can tolerate. His ways of dealing with new stimuli certainly should tell us something about how he will handle interpretations given him. Ultimately, it should reveal something about the nature of his object relations. After all, another person represents a mighty powerful stimulus which one can ward off, whose intrusion one might fear, and whose real and separate existence one might be oblivious to, by whose presence one may feel overwhelmed, or to whom one may feel open and receptive.

In making a developmental diagnosis it is also instructive to gauge at what level of differentiation a child

operates with respect to any particular ego function. For example, does he produce intact material which is, however, primitive and lacking any kind of elaboration? Or does he offer highly differentiated material, and if so, are his synthetic functions equal to handling this degree of differentiation? I recall a case in point in which an evaluation was done at the beginning of treatment and another a year later. This particular girl's difficulties had, in all likelihood, an organic substrate. Her first record was primitive, simple and in many ways defective. With growth and treatment, her second evaluation revealed infinitely more complex, highly elaborated material. However, since this girl's synthetic functions were impaired she also looked a lot "sicker" on the second evaluation. Her drawings of a person, for example, had changed from being globally intact, primitive blobs to attempts to account for the finer physical characteristics such as features, fingers, etc. The result was a grossly bizarre, distorted, though detailed rendition of the human body. One can readily see that the variable of "level of ego differentiation" is sensitive in picking up change, and thus is a variable that can profitably be used in estimating progress in therapy.

II. *Defense Diagnosis.* While the defenses are part of the ego, they nevertheless warrant separate discussion. A developmental diagnosis is incomplete without an assessment of the defense structure. Clinically we know that most defensive efforts have an adaptive function. In evaluating the adjustment of a child we are called upon to judge whether the pathogenic aspects of a defense outweigh the adaptive ones. In relatively healthy children this may be a difficult task. One way of differentiating is to gauge how fluid a child's defense structure still is or how set and characterologically fixed his ways of coping have already become. In the way he responds to the tests and the testing situation

we can learn how invariably he uses one or the other defense or how versatile he is in his defensive maneuvers. We may also get a clue as to what his preferred defenses might be. Our expectations of how a child will respond to treatment will differ if he is primarily a denier, or an externalizer, an isolator or a projector, a counterphobe or a complier. The latter, who is relatively easy to spot through tests, poses a particular treatment problem: in the long run the complier may be most resistant to genuine gains, for to please us he will show all sorts of changes.

Test records vary of course in the extent to which they will reveal a child's defensive structure. The reason for this lies partly in the test situation, and partly in the nature of the defenses themselves (e.g., compulsivity is much more easily witnessed than repression). The test situation may evoke certain types of defenses (and defensiveness) more readily than other situations, and this may become a problem for differentiating between what is centrally and what is peripherally important. Also, there is always the question whether the observational sample gathered through tests is large enough to allow expression of all the different defenses which are at the child's disposal. In addition, there is the problem described by Anna Freud (1937), namely, that successful defense is carried out silently and invisibly, with the ego knowing nothing of the rejected impulse nor of its defensive maneuvers.

III. Considerations of Superego Pressures, Self Image and Ego Ideal. Interesting clues regarding these can be gained through the way a child takes the tests. Some children are so humiliated by the thought of failing they will not even tackle the task. They dare not expose themselves to the possibility of failure. This may lead from manifestations of too much conscience to too little conscience: the child in his refusal to face failure may

save himself a great deal of effort. Or, he may be self-critical of his efforts for a variety of face-saving reasons (see Korner (1961)) such as, for instance, the hope that this will ward off criticism which he anticipates. There are other maneuvers geared to an introjected critical audience. For example, some children introduce most of their statements with a congenial "you know" or "of course" as if to deny the wish that an idea might be uniquely theirs, hoping to ward off through this renunciation expected competitiveness on the part of the examiner. Altogether, the child's use of the examiner in the testing situation is full of significant clues. Does the child appoint him as judge? Or does he already have well established internalized critical functions which permit him a realistic appreciation of his successes and failures? Does he disown responsibility, blaming the test or the tester for his shortcomings? Do his remarks suggest he wished he were different and in what way? Does he speak with his own voice in making these remarks or can one sense in the shifts in his voice that internalized objects are speaking for him? It is obvious that much can be learned from these utterances about the maturity and severity of his conscience, about his self-image and possibly his ego ideal. Needless to say, elements of these also transpire in the way he deals with his TAT characters.

IV. Evaluation of Instinctual Components. Instinctual components, unresolved libidinal conflicts and fixations express themselves in a variety of ways on tests. First of all, instinctual components and psychosexual fixation points give a very specific cast to a child's mode of thought and to his work approach. For example, unresolved oral fixations drive some children to bite off and to digest only little pieces of a problem, rendering them incapable of taking in its totality. Their approach is impatient, impulsive and voracious. They are the poor delayers, and their development

is still overshadowed by the earliest oral mode described by Erikson (1950). Then there are those who, after taking in a problem, chew it over *ad infinitum*, a process which reflects remnants of either oral or anal retentiveness. Then there are those who, in response to any question, will spew forth a mass of unrelated thoughts reminiscent of the anal-eliminative mode. Children who have not settled autonomy issues may suffer severe cognitive inhibitions: they may simply not be in a position to take in instructions at the time they are given to them. This is true also of children who for reasons of castration anxiety fear passivity. Certain tasks require entering into and analyzing parts of the problem. This takes an active intrusiveness with which some children with phallic inhibitions cannot cope. They tend to stay strictly on the surface of the problem, trying to solve it by global approximations. Even sexual identification can be reflected in thought form. Male identified children of either sex frequently think with incisiveness and precision whereas children who defend themselves against such an identification may show this in vague and fuzzy thinking.

Secondly, instinctual components may be inferred through drive-dominated thinking. A good illustration of this type of thought is the answer of the son of a retired Army man who interpreted the proverb: "He who would eat the kernel must crack the nut" as "People aren't usually cannibals." Defensive decompensations can readily be witnessed in the breakthrough of such drive-dominated material. In Hartmann, Kris and Loewenstein's terms (1949) such children are lacking in neutralized energy with which to tackle realistically intellectual tasks. Sometimes these children hypercathect situations with aggressive or sexual implications, seeing these more clearly than any other situations. More often, these children's understanding becomes fuzzy whenever a problem approaches forbidden im-

pulses. Occasionally, one encounters repetitions of completely innocuous remarks; on careful probing one finds a wealth of affective material hidden under these.

Finally, and the most obvious way in which instinctual components reveal themselves is in the content of the child's answers. Projective techniques, in particular, can reveal a great deal about a child's subjective reality, the nature of his instinctual conflicts and the stage of his psychosexual development. These tests very readily bring out conflicts of sexual identity. They also highlight age confusion which often is a particular problem for the child with unresolved oedipal strivings. They frequently see children on the TAT as grownups or older people as young ones; the distortions often bring about wishfulfillment of an oedipal fantasy.

Of particular interest regarding the breakthrough of instinctual material are the records of young adults who have a history of consummated incest in early childhood. I have had occasion to examine eight such adults, and to consult on many more, all women who had incestuous relations with either their fathers or stepfathers. By virtue of being clinic patients they do not necessarily reflect the adjustment of women with incest experiences in general. Although clinically nonpsychotic, all but one of the eight patients presented Rorschachs filled with primary-process type thinking, particularly contaminations, condensations and displacements. There was multiple evidence in these records that unrepressed instinctual manifestations had a disruptive effect on both ego organization and defensive structure. In focal areas Rorschach imagery suggested that these patients all had a blurred reality sense. This was particularly true in the sexual area in which these patients did not clearly differentiate male and female. Altogether, judging from their contaminatory tendencies these patients had

rather fluid ego boundaries. Some of their TAT's suggested age confusion. The records of these patients reflected instinctual pressures breaking through at all levels: in mode of thought, in drive-dominated thinking and in content. This is a good example of how tests can reveal end products but not their etiology. The latter we can fathom only through what psychoanalysis has taught us. We know from that source that fulfillment of what are usually unconscious wishes (such as incest) may result in confusion between reality and fantasy. We also know that such fulfillment may make repression impossible. Thus, what is usually unconscious remains conscious, interfering in terms of its affect load with proper ego functioning. As a result we may see all kinds of manifestations of primary process thinking. Also, considering the effect of massive instinctual stimulation on an immature organism which does not as yet possess mature sexual discharge channels, one may expect perceptual distortions, and especially confusion in body imagery. Furthermore, by taking the place of a grownup, the mother, role and age confusion are a likely development. This is also true of persistence in omnipotent thinking with the power to expose and to depose which these children have actually experienced.

V. Genetic Considerations. Since we witness through psychological tests largely end products of highly elaborated, over-determined genetic events, we are up against great odds in trying to make accurate genetic reconstructions. The best we can hope for are hunches and approximations. Even though these represent only parts of what were probably true events, they are nevertheless worthwhile as working hypotheses to be revised and enlarged as further material unfolds.

While test results rarely yield genetic answers they can be used most profitably in raising such questions.

mental unevenness we would want to inquire into the history of this. We would ask questions like: Were there any reversals in the natural chronology of the maturational steps during infancy? Were there variations in the receptivity of the various sensory modalities favoring the development of certain functions, hindering others? Or have, perhaps, certain functions not been allowed to develop because of a symbiotic tie to a parent? Or let us take the example of an intensely hyperactive child. The origin of this child's hyperactivity has vast diagnostic and prognostic ramifications. Has this child even as an infant needed this type of tension discharge? Or did his difficulties start when he began to walk and get into things which put him at odds with a mother who, up to this point, could meet his needs, but who, for reasons of her own, could not tolerate his locomotor intrusiveness? Or does he have a history of immobilization due to illness or injury just at the time when he learned to walk? Or is his hyperactivity a defense against passivity, or a masturbatory equivalent? Or do we find that hyperactivity is part of the family pattern in which he can find several models of identification? Test observations thus can be used to raise genetic questions. They can pinpoint gaps in the developmental history which can be meaningfully filled. They thus can aid in the diagnostic process, which a treatment situation is in part also by raising questions important to the ultimate reconstruction of the case.

SUMMARY

One of our best criteria for treatment of childhood disorders is the arrest of any aspect of the developmental process. This criterion helps view isolated symptoms in perspective, allows, without undue concern, for transient disturbances attendant to the mastery of new maturational steps and avoids stereotypic standards of psychological health. An evaluation of the child's developmental

process requires an appraisal of the intactness of the child's libidinal and ego development. Psychological tests can help spell out the level of maturity, regressions, fixations and progressions both in libidinal and ego terms. They are especially suited for revealing the nature and the level of differentiation of specific ego characteristics. Potentially, they also are helpful in gauging the fluidity of the defense structure, the maturity and severity of superego pressures and the nature of the child's self-image and ego ideal. Level of libidinal development is reflected in mode of thought, breakthrough of drive-dominated thinking and in content. Perceptual style, in particular, takes on characteristics of Erikson's "organ modes," thus revealing the status of a child's psychosexual development. Tests primarily reveal end products of development; they do not clearly differentiate between the influence of innate ego variations and acquired ego characteristics nor between the effects of libidinal endowment and childrearing methods on psychosexual development. Tests then, at best, only furnish genetic and etiological hunches. They do, by contrast, raise a great many genetic questions by revealing areas of developmental unevenness. By assessing the relative maturity, flexibility and capacity to progress of a variety of dimensions expressive of libidinal and ego development, psychological tests are a valuable aid in making a developmental diagnosis.

REFERENCES

- Anthony, E. J. The significance of Jean Piaget for child psychiatry. *Brit. J. Med. Psychol.* 1956, 29, 20-34.
- Anthony, E. J. The system makers: Piaget and Freud. Symposium on the contribution of current theories to an understanding of child development. *Brit. J. Med. Psychol.* 1957, 30, 255-269.
- Benjamin, J. D. The innate and the experiential in development. In: *Lectures on experiential psychiatry*. Pittsburgh: Pittsburgh Univ. Press, 1961, 19-42.
- Erikson, E. H. The theory of infantile sexuality. In: *Childhood and society*. N.Y.: Norton & Co., Inc. 1950, 44-92.
- Escalona, Sibylle & Heider, Grace M. *Prediction and outcome*. N.Y.: Basic Books, 1959, 73-81.
- Freud, Anna. *The ego and the mechanisms of defense*. N.Y.: Hogarth Press, 1937.
- Freud, Anna. Indications for child analysis. *The psychoanalytic study of the child*. 1945, 1, 127-149.
- Freud, S. The predisposition of obsessional neurosis. A contribution to the problem of the option of neurosis. *Collected papers, II*. London: Hogarth Press, 1946, 122-132.
- Freud, S. Three essays on the theory of sexuality. II. Infantile sexuality. *The standard edition of the complete psychological works of S. Freud, VII*, London: Hogarth Press, 1953, 173-207.
- Fries, Margaret & Wolff, P. J. Some basic differences in newborn infants during the lying in period. N.Y. Univ. Film Library.
- Gardner, R., Holzman, P. S., Klein, G. S., Linton, H. & Spence, D. P., Cognitive controls. *Psychological Issues*, 1959, 1, 4.
- Gesell, A. & Amatruda, C. S. *Developmental diagnosis*. N.Y.: P. H. Hoeber, Inc., 1947.
- Hartmann, H., Kris, E. & Loewenstein, R. M. Notes on the theory of aggression. *The psychoanalytic study of the child*, 1949, III/IV, 9-36.
- Hartmann, H. Comments on the psychoanalytic theory of the ego. *The psychoanalytic study of the child*, 1950, V, 74-96.
- Korner, Anneliese F. Limitations of projective techniques: Apparent and real. *J. Proj. Tech.* 1956, 20, 1, 42-47.
- Korner, Anneliese F. The parent takes the blame. *Social Casework* 1961, XLII, 4, 339-342.
- Piaget, J. *The child's conception of physical causality*. N.Y.: Harcourt Brace & Co., 1930.
- Piaget, J. *The origins of intelligence in children*. N.Y.: International Universities Press, Inc., 1952.
- Terman, L. M. & Merrill, Maud A. *Measuring intelligence*. Boston: Houghton Mifflin Co. 1937.
- Werner, H. The concept of development from a comparative and organismic point of view. In: *The concept of development*, edited by Harris, D. B. Minneapolis: Univ. of Minn. Press, 1947, 125-148.
- Wolff, P. The developmental psychologies of Jean Piaget and psychoanalysis. *Psychological Issues*, 1960, II, 1.

Received October 31, 1961

The Prediction of TAT Structure as a Test of Rorschach's Experience-Balance¹

JAMES O. PALMER AND BILLIE J. LUSTGARTEN

The Neuropsychiatric Institute, University of California Medical Center, Los Angeles, California

Various students of personality have hypothesized and partly demonstrated that certain selective perceptual processes appear to underlie and circumscribe most personal and social attitudes. The attempt to predict such attitudes, and thereby emotional functioning and personality structure, through observing and measuring certain perceptual selection has been the chief unique contribution of Hermann Rorschach (1942). The interrelationships between movement, color and form—the classic variables in studies of perception—are the constituents of his central hypothetical construct: the Experience Balance. However, the Experience Balance (hereinafter referred to as the E-B), though known by name even to the casual reader of Rorschach, is not commonly understood in terms of perception, being perhaps more frequently associated with the less perceptually defined “introversion-extroversion” concepts of Jung. Restated briefly, the essence of the E-B is that it refers, in Rorschach's words, to methods of “experiencing” (in current terminology: “perceiving”) rather than “living” (for which read “behaving”). Thus the subjective projection of movement as a method of organizing a static form is, in Rorschach Theory, analogous to the use of pre-formed subjective attitudes—as the basis of judgment for social and emotional reactions. The term “intratension” is applied when such subjective attitudes are the most frequent bias. Similarly, extratension, the predominant use of bright color as a response determinant, theoretically predicts

not emotional environmental interaction per se but *dependence upon emotional cues* from the environment. According to the E-B concept, both the intratensive and the extratensive person may behave in a socially expansive or a withdrawn fashion but the basis for judging the environment which determines their reaction is assumed to be different. It should be kept in mind that Rorschach also posited two other perceptual modes: the “ambiequal” type in which both external and internal cues (color and movement) are used, and the “coarctate” type which ignores both cues.

Despite the fascinating implications of Rorschach's hypotheses for the study of both perception and personality, these hypotheses have until recently been disregarded by (or unknown to) experimental students. Whether because of its abstruse relationship to clinical practice or the absence of experimental support, the E-B has been given little more than lip-service in the wide-spread clinical use of the Rorschach Test. In the last decade, however, several investigators have made at least tentative and exploratory attempts to follow Rorschach's lead. The aim of these studies has not been so much to sharpen clinical test interpretation as to broaden the basis of understanding of the relationships between personality and perception. Outstanding have been the studies of Meltzoff (1953), Singer (1952) and their colleagues wherein they repeatedly demonstrate the relationship between inhibition process, motor delay and perception of movement in inkblots. Bieri and his colleagues (1956, 1957) found

¹This study was made possible by a grant from the U.S. Public Health Service, M.

lated to "experience type." Verbal and fantasy correlates of inkblot perceptual determinants have been demonstrated by Gellerman and Sloan (1951), by Carp (1950) and by Shatin (1955, 1958).

In our own prior studies (Palmer, 1955, 1956 & 1957) correlates of the E-B were found in such widely diverse behavior as problem-solving on I.Q. tests, patterns of responses to the MMPI and reactions to frustration (on Rosenzweig's Picture Frustration Study). However, the interpretation of the behavior of Rorschach-typed subjects on these other tests was admittedly post facto. For example, a sample of a clinic population was divided into four sub-samples on the basis of E-B scores, an item-analysis of the MMPI responses of each E-B group was conducted, and the results were interpreted afterwards in the light of Rorschach's E-B hypotheses. Thus, in these intentionally "exploratory" studies, no specific predictions were attempted, and cross-validation is lacking. Rather, it was hoped that from such explorations specific hypotheses might appear which would permit and promote prediction.

Skipping over, at least temporarily, the labors of cross-validation of these prior studies, the present study is intended instead as a predictive test of the E-B hypotheses. As an object of prediction, the structure of responses to the TAT offered several possibilities. Prediction of the structure or form of such fantasy behavior seemed more possible as the E-B hypotheses refer to determinants of structure of perception rather than to ideational content. Furthermore, TAT structure seems *prima facie* to have elements in common with the E-B. The E-B purports to classify people according to the extent that they use personal fantasy as a basis of perception; reliance on fantasy seems to play a part in the subjects' ability to carry out Murray's (1943) instructions for his TAT, i.e., to construct a plot of

three elements: a present, past and future. Murray also specified that the characters should be described and that the "feelings" of the characters be indicated; dependence upon external or descriptive characteristics of the environment is allegedly the central feature of extratension, while awareness of intrapersonal feelings is supposed to be the hallmark of the intratensive person. The possibility of a relationship between this aspect of TAT structure and the E-B has already been given support by Hays, Gellerman and Sloan (1951) who, using a verb-adjective quotient as a TAT score, successfully predicted greater use of verbs by intratensives and adjectives by extratensives. Last but not least, the central feature of Murray's personality theory (Murray et al. 1938) is directly analogous to the E-B dichotomy, i.e., the contrast between the person's use of his internal *needs* as a basis for his perceptions and behavior as opposed to the *press* of the external environment upon his judgments and acts.

The following hypotheses were thus constructed for testing:

1. If the intratensive type ($M > \text{Sum } C$) perceive on the basis of internal stimuli such persons should
 - a. yield more complete plots than either the extratensive or
 - b. the coartate group.
 - c. be more concerned with the feelings of the persons in their stories than the extratensives and
 - d. the coartates.
 - e. be more concerned with "needs" in the motivations of heroes in their stories than the extratensives or
 - f. the coartates.
2. If the extratensive type ($\text{Sum } C > M$) are dependent upon environmental stimulation, such persons should:
 - a. be more concerned with de-

- scriptive characterization of heroes in their stories than the intratensives.
- b. more concerned with environmental press on their heroes than either the intratensive or c. the coartates.
3. If the coartate type ($M = \text{Sum C}$ and both below their median scores) avoids the use of emotional cues in his perceptions, such a person should:
 - a. yield only bare plots, scarcely more than descriptions of the TAT cards.
 - b. be less concerned than any of the other Rorschach types with either the feelings of or
 - c. characterization of story heroes.
 - d. fail to identify clearly the motivations of their heroes, either internal needs or
 - e. environmental press.
 4. If the ambi-equal type ($M = \text{sum C}$ and both above median scores) utilize both intra- and extra-personal emotional cues in perception, then such persons should:
 - a. show generally better TAT plot construction than any of the other E-B types
 - b. be equally concerned with both the hero's character and feelings.
 - c. be equally concerned with needs and press in the hero's motivations.

PROCEDURE

Sample: All cases which contained at least 10 TAT stories and Rorschachs of 10 responses or more were selected from the population of all adult, white out-patients examined at the Washington University Medical School, St. Louis, (1949-1952) and at the University of California Medical Center, Los Angeles (1955-1959). The sample was further restricted to

one of the four Rorschach E-B categories, according to the scoring criteria described below. The 79 men and 45 women who met these criteria ranged in age from 18 to 61, mean age 31.2. Diagnostically, most of these 124 patients fell under the entries of "neurotic reaction" or "personality trait disturbance" (most frequently "passive-aggressive"). Wechsler Intelligence Scale I.Q.'s (WB-I or WAIS), available on 55 of these patients, ranged from 69 to 143, with a mean I.Q. of 110.5. As in the sample studied in 1953, the A group (Mean I.Q. 116.27) and the I group (Mean I.Q. 110.08) appear generally to function with greater intellectual efficiency than either the E group (Mean I.Q. 108.94) or the C group (Mean I.Q. 106.45), although these differences are not statistically significant on an analysis of variance ($F = 0.889$). Results of other I.Q. tests (mostly Shipley-Hartford) likewise indicated that our sample is of above average to superior intellect.

E-B Classification: Since the tests were administered by various staff members and interns, all Rorschach protocols were rescored, using the criteria described in the 1953 study. Rorschachs with incomplete inquiries were per force eliminated. The EB classification of each case was psychometrically determined, using the formula derived from Hertz (1943):

$$EB = 100. \left[\frac{M - \text{Sum C}}{R} \right] + 100. \text{ This}$$

formula has the advantage of turning the ratio into a score. Since this score is a percent of the total number of responses, R , in effect, is held constant (at least in the range of 10 - 40 responses, which covers most of our cases). These scores were normally distributed, with a median of 98; cases below Q_1 , 89, were deemed extratensive and those above Q_3 , 106, were called intratensive. Within the range 90 - 105, the dilated or ambi-equal group was distinguished from the constricted or coartate cases by

separate ratios M/R and Sum C/R: if both ratios were above the median, the case was considered ambi-equal. The balance of cases in this center range of the distribution was considered coartate. Retrospectively it became apparent that these cutting points for intratension and extratension were lacking in rigor; too many cases were within one or two responses of falling in one group or another. In the present study a cut-off point of 2 SD was used, such that only records with scores of 83 or less were deemed extratensive and only those with scores of 112 or more called intratensive. More rigorous definitions were also given the coartate and ambi-equal groups: Coartation was limited to those cases in which the ratios M/R and Sum C/R were *both* below their respective medians. As before, only those cases in which *both* scores were above these medians were dubbed ambi-equal.

Administration and Scoring of the TAT: The TAT instructions and method of administration are presumed to have been that recommended by Murray, though admittedly there was no control over this factor. Examination of the recorded stories (all of which were close to verbatim) indicated that most examiners tended to encourage the subject to tell as complete stories as possible on the first several cards; fewer notes indicating questions or encouragement appear in the latter half of the test. The stories ranged in length from 30 to 615 words, with a mean of 105 words, $SD = 60.2$.

Each of the 1336 TAT stories was scored on five seven-point scales, as follows:

PC, "Plot-construction": the degree to which the subject used Murray's three elements, "past, present, future", ranging from 0 "no plot, description only" to 6 "all three elements."

C, "Characterization": the number of adjectives or adjectival phrases de-

scribing persons or situations (other than simple identifying adjectives, such as "young" or "old" where different ages are suggested by the picture). Stories were scored from 0 (no such adjectives) to 6 (six or more).

F, "Feelings": the number of verbs, verb forms or phrases indicating the thoughts or feelings of a person (but not simple physical actions), again ranging from none (0) to six or more (score 6).

N, "Needs": the number of expressions of any kind of motivation attributed to any person in the story, i.e., arising from the person as contrasted with stimulation from the environment, again scored from 0 to 6.

P, "Press": the number of motivations attributed to stimulation from the environment, scored 0 to 6.

Although no predictions were made regarding TAT content, the stories were also rated on an additional variable, *V, "Violence"*. On this scale, stories were rated from 0 (no mention of any aggressiveness) to 6 (mention of homicide).

After a pre-trial scoring on a separate sample, during which each variable was scored separately, all variables were scored at the same reading of each story. However, the variables were scored by picture number, never by subject, so as to avoid subject halo. The scores on the 10 stories were then added to obtain a total score for every subject on each variable, (pro-rated when more than 10 pictures were administered). All results refer to these total scores.

Scoring Reliability: Reliability of the Rorschach scoring was not rechecked in this study, since it was based on fairly objective criteria and has been shown to be fairly high (Ramzy and Pickard 1949). Of greater concern was the reliability of the TAT ratings which, although defined as carefully as possible, and pre-tested on separate trial sample, often seemed difficult to apply. Table I

shows intercorrelations testing three kinds of reliability: interjudge (between the authors) on 20 cases, intra-judge (20 cases re-scored by the junior author) and split half reliability (using the first five cards versus the second five, according to the numbering of the cards by Murray). Apparently the junior author who made the final ratings was fairly consistent in her scoring (this intrajudge check was made after the interjudge reliability showed some inconsistencies between judges and the scoring criteria had been sharpened). However, the split-half reliabilities show that the judges were laboring with a test on which subjects are rather inconsistent from card to card in indicating the character and feelings of their heroes, and responding almost completely random in mentioning motivations.

Since the 10-13 pictures selected by the clinicians varied from case to case, the effect of card selection on the variables under consideration was examined. A study of "card-pull" on a separate sample indicated that each of these scores varied widely from picture to picture. For each score, two groups of pictures were distinguished: those with high scores (T scores of 60 or above on the variable under consideration) and those with low (T scores of 40 or below). The frequencies with which each of these groups of pictures were administered to the

four Rorschach types were compared. Since no differences among E-B types were found for any set of pictures, it may be assumed that card-pull did not significantly affect the findings.

Another question of consistency or "reliability" might be phrased, "how distinct were the TAT variables from one another?"—particularly since they were scored simultaneously on each story. The low positive intercorrelations between variables shown in Table II indicates that for the most part the variables were uncontaminated. Since the judgments of "feelings" were seldom based on the same wording used to score "press", this single high correlation appears to be a chance result. The next highest r of .51 between plot-construction and need suggests that more needs are indicated in the more complete stories. Otherwise it is interesting to note that neither plot-construction (story completeness) nor story length (word count) have more than very slight relationship to the other variables—or to one another.

RESULTS

The means and mean differences between each pair of Rorschach groups on each of the TAT variables are shown in Table III. Seven of the fifteen hypotheses were supported by the significant differences at the .05 level or better. In six other instances the differences were in the expected

TABLE I—Reliability of TAT Ratings

Type of Reliability	N	Plot Constr.	Characterization	Feelings	Needs	Press	Violence
Inter-judge	20	.92	.46	.66	.55	.57	.75
Intra-judge	20	.88	.86	.89	.60	.62	.66
Split-Half	124	.61	.57	.47	.16	.17	.56

TABLE II—Intercorrelations Between TAT Variables

	C	F	N	P	V	L*
Plot Construction	.37	.16	.51	.22	.29	.05
Characterization		.25	.40	.02	.13	.28
Feelings			.44	.88	.00	.33 ¹
Needs				.11	.05	.20
Press					.16	.05
Violence						.02

* Story length as measured by word-count

TABLE III—TAT Means and Mean Differences Between Experience Balance Groups
Rorschach Group Means

TAT Variable	I-E					I-C					I-A					E-C					E-A					C-A							
	I (35)	E (35)	C (22)	A (37)	e	I-E	t	d	e	o	I-C	t	p	e	o	I-A	t	p	e	o	E-C	t	p	e	o	E-A	t	p	e	o	C-A	t	p
	36.71	29.50	30.50	33.39	>	>	>	2.85	>	>	>	2.45	.005	—	—	>	>	ns	—	>	>	—	1.71	.05	>	>	>	>	>	>	1.34	.10	
	24.23	21.44	16.64	23.54	>	>	>	1.44	>	>	>	2.34	.025	—	—	>	>	ns	—	>	>	—	>	>	>	>	>	>	>	>	>	2.60	.01
	11.63	9.67	8.36	11.76	>	>	>	1.10	>	>	>	1.19	.10	—	—	>	>	ns	—	>	>	—	>	>	>	>	>	>	>	>	>	2.18	.01
	8.43	7.27	6.23	8.11	>	>	>	ns	>	>	>	2.54	.025	—	—	>	>	ns	—	>	>	—	>	>	>	>	>	>	>	>	>	2.47	.01
	5.89	5.37	4.67	6.41	>	>	>	ns	>	>	>	2.04	.050	—	—	>	>	ns	—	>	>	—	>	>	>	>	>	>	>	>	>	2.76	.01
	14.57	13.57	16.64	14.59	>	>	>	—	>	>	>	3.01	.005	—	—	>	>	ns	—	>	>	—	>	>	>	>	>	>	>	>	>	2.09	.05

LEGEND

() Contain N's

e shows direction of predicted differences; — indicates no prediction made; o shows direction of obtained differences. Single tailed tests made on all predicted differences; all others are two tailed tests. Only significant t's are entered in the table.
TAT variables: PC, Plot Completion; C, Characterization; F, Feelings; N, Needs; P, Press; V, Violence.
Rorschach groups: I, Intratensives; E, Extratensives; C, Coartates; A, Ambiequals.

direction, though lacking acceptable statistical significance. In two instances the differences are counter to those predicted, but not significantly so. There were no statistically significant differences contradicting these predictions. When these data were broken down by sex, the results for the males were grossly the same as for the total group. The differences among the female Rorschach types seemed less sharp and consistent; however, the female sample was relatively small.

Before examining the specific hypotheses in the light of these results, it should be noted that two groups, I and A, yielded generally higher mean scores on nearly all variables. Possibly this was because these groups told the more complete stories, as indicated by their higher mean PC scores; however, PC is not highly correlated with any other variable, with the exception of needs. Or it may be a factor of intelligence, for, although no overall difference in I.Q. among the four Rorschach groups was found in this sample, on a larger sample (including some subjects of the present sample) it was established that the I and A groups have higher mean I.Q.'s (Palmer, 1953).

Examining the specific hypotheses, higher scores were predicted on most all TAT scales for A and I, with the exception of the prediction that E groups would be higher than I on characterization and needs. Coartates were predicted to be low or lowest on most of these scales. Thus as foreseen in hypothesis 1., the intratensive group told more complete stories than (a)² the extratensives or (b) than the coartates, as measured by higher PC scores. The intratensives, as expected, had (c) a higher mean score on feelings than the extratensives and (d) similarly emphasized feelings more than the coartate group. Intratensives do not, contrary to the pre-

²These letters in () refer to the sub-sections of the hypothesis under consideration.

diction, (e) emphasize internal needs any more than do the extratensives, but are (f) more aware of needs than the coartate group.

In contrast, the extratensive group failed to show the differences expected in hypothesis 2. They were (a) no more concerned with the external characteristics of their heroes nor (b and c) did they use environmental press as a motivation more than any other group; indeed, their obverse, the intratensives, yielded slightly higher mean scores on both these TAT variables. The coartate group, as predicted in hypotheses 3 (a) also told "coartate" stories (low PC scores), were unconcerned with either the feelings or the character of their heroes (b and c) and failed to specify motivations, either need (d) or press (e). Concerning hypotheses 4, the ambiequals did not tell more complete stories than other groups, contrary to our prediction (a), but were equally concerned with both feelings and characterization and with need and press as predicted (b and c).

However, these latter predictions concerning the ambiequal and coartate groups are phrased as comparisons between TAT scores within the respective Rorschach group rather than between groups. When mean T-scores, shown in Table IV, are examined, some further confirmation of the hypotheses is provided. For the ambiequal group, each of the TAT T-scores is approximately at the mean, with little or no difference among them. Just as the ambiequal person pays attention to both movement and color on the Rorschach, he utilizes equally each structural element of the TAT. On the other hand, the mean TAT T-score "pro-

file" of the coartate group is consistently below the mean (except for the unpredicted "Violence" score), with the low point on the "Needs" scale. Thus, the failure to recognize motivation on the TAT seems to characterize the person who also fails to use either color or movement on the Rorschach.

DISCUSSION

Rorschach's general hypothesis concerning introversion appears to be given some further support by these results. Persons who project human movement onto his ink-blots do, as he suggested, tend to be more imaginative (in that they tell more complete plots) and at least tend to recognize internalized feelings and needs. However, they are not completely unaware of external factors, at least on this fantasy level. Why such an esoteric behavior as seeing M on inkblots should be grossly related to fantasy behavior (or to intelligence or patterns of MMPI attitudes) is, for these authors, a fascinating but as yet unfathomed inquiry. The best lead to an answer seems promised by the "sensory-tonic" theory in which M is considered a product of the delay or inhibiting function of the ego (Meltzoff, Singer and Korchin 1953), especially as related to inhibition of motor activity.

The failure to substantiate the hypotheses concerning extratensive behavior, e.g., emphasis on external characteristics and motivations, may be due in part to the nature of the criterion. Extratensives seem relatively inadequate at fantasy production in general. However, this finding does indicate that conceptualization about these dimensions must be more

TABLE IV—Mean T-Scores on TAT Variables for each Rorschach Type

Rorschach Group	TAT Variables				
	PC	C	F	N	P
A	50.80	51.25	51.99	53.66	52.85
C	47.88	45.65	46.13	38.38	45.86
E	46.87	50.54	48.38	46.83	48.67
I	54.55	51.81	50.03	50.00	50.76
					50.41

sharply delineated. If extratension is a definite personality dimension, and if it is measured by "sum color" on the Rorschach, then it is limited at least to a non-fantasy level. Perhaps extratension would become more demonstrable in interpersonal relationships or under other environmental stress. On the other hand, would intratension fail to show up under externalized situations?

Coartation and ambiequality appear to be definite dimensions underlying perceptual selection in fantasy structure. These findings are in accord with Carp's (1950) study of "constriction" on projective techniques and Shatin's several studies (1953, 1955 and 1958) on the constriction-dilation dimension in Rorschach and TAT. The coartate group's significantly higher mean score on the TAT "violence" rating suggests that coartation of the use of emotional cues in perception may not mean a coartation of emotional ideation. Similar concern with "violent" and hostile behavior appears in the MMPI responses of the coartate group (Palmer 1956). The occurrence of this phenomenon on both the fantasy and subjective attitude level lends further weight to the interpretation that when both internal and external cues of emotion are disregarded, the individual has little basis for judging and controlling either his environment or himself. The result may be a lower threshold of frustration and more diffuse and labile reactivity. However, in standard Rorschach interpretation, coartation has been interpreted as meaning an absence of emotionality while lability and diffuseness of reactivity has been attributed to extratension. Yet these results show extratensives to be lowest on the TAT violence scale, and this group was shown to have emotionally conservative, even rigid, reactionary attitudes on the MMPI (Palmer 1956). In our opinion the usual interpretation of color responses on the Rorschach disregards

Rorschach's distinction between perception and response. Evidently extratension, if it does mean awareness of environmental stimulation, does not necessarily mean emotional reactivity.

SUMMARY

Hypotheses, derived from prior exploratory studies, were constructed predicting the structure of TAT fantasy from the Rorschach E-B score. Out of a population of adult clinic patients to whom both the Rorschach and the TAT have been administered, a sample of 79 men and 45 women was selected such that each of the four E-B groups were equally represented. These E-B groups were psychometrically determined, using sharply delineated cutting-points. The TAT stories were independently rated as to "plot construction", "characterization", "feelings", "press", "needs", and "violence". Interjudge reliabilities ranged from .46 to .92 and intercard reliability from .16 to .61. Nevertheless, differences between the E-B groups on the cumulative TAT ratings supported seven of the fifteen predictions. However, Rorschach extratensives do not emphasize the external personality characteristics of their TAT heroes nor environmental press as motivation, contrary to our expectations, but as this E-B group has difficulty in general in creating fantasies (as might be expected) the TAT is probably not an adequate vehicle for demonstrating the perceptual attitudes of the extratensive. On the other hand, intratensives, who did prove to have the more elaborate and organized fantasies, in general emphasized the internal feelings and motivations of their heroes; predicted differences between intratensives and coartates were statistically clear-cut but these variables did not distinguish the intratensives from the extratensives as was hoped. As was predicted, the ambiequals tend to ascribe both internal and external characteristics and motivations to their heroes

while coartates yielded the least complete stories, barren of motivation, with ill-defined heroes. No predictions were made regarding TAT content but ratings of "violence" turned out to be much higher on the coartate group than for any other E-B type, while the extratensives tended to express the least "violence", a finding consistent with prior studies (Palmer, 1956).

On the whole the prediction of perceptual behavior in such nebulous and errant material as TAT fantasy proved an exacting test. The possibility that perception of movement and color is predictive of perceptual attitudes in other behavior thus remains encouraging.

REFERENCES

- Carp, Frances. Psychological constriction on several projective techniques. *J. consult. Psychol.*, 1950, 14, 268-275.
- Hays, W., Gellerman, S. & Sloan, W. A study of the relationships between the verb-adjective quotient and The Rorschach Experience Balance. *J. clin. Psychol.*, 1951, 7, 224-227.
- Hertz, Marguerite. Personality patterns in adolescence as portrayed by the Rorschach Ink Blot Method: III The "Erlebnestypus" (a normative study). *J. genet. Psychol.*, 1943, 28, 225-276.
- Meltzoff, J., Singer, J. L. and Korchin, S. J. Motor inhibition and Rorschach movement responses: a test of the Sensory-Tonic theory. *J. Pers.*, 1953, 21, 400-410.
- Murray, Henry et al. *Explorations in Personality*. 1938. New York: Oxford Univ. Press. pp i-xii, 1-761.
- Murray, Henry. *Thematic Apperception Test Manual*. 1943. Cambridge: Harvard Univ. Press.
- Palmer, J. O. Rorschach's Experience Balance: The concept, general population characteristics and intellectual correlations. *J. proj. Tech.*, 1955, 19, 138-145.
- Palmer, J. O. Attitudinal correlates of Rorschach's Experience Balance. *J. proj. Tech.*, 1956, 20, 207-211.
- Palmer, J. O. Some relationships between Rorschach's Experience Balance and Rosenzweig's Frustration Aggression patterns. *J. Proj. Tech.*, 1957, 21, 137-141.
- Ramzy, I. & Pickard, P. M. A study in the reliability of scoring the Rorschach Ink-Blot Test. *J. gen. Psychol.*, 1949, 40, 3-10.
- Shatin, Leo. Rorschach adjustment and the Thematic Apperception Test. *J. proj. Tech.*, 1953, 17, 92-101.
- Shatin, Leo. Relationships between the Rorschach and the Thematic Apperception Test. *J. proj. Tech.*, 1955, 19, 317-331.
- Shatin, Leo. The constriction - dilation dimension in Rorschach and TAT. *J. clin. psychol.*, 1958, 14, 150-154.

Received September 29, 1961.

The Application of a New Matching Technique

GEORGE H. WEINBERG, FRITZ A. FLUCKIGER AND CLARENCE A. TRIPP
The Handwriting Institute, Inc., New York City

The matching technique has been used often in connection with investigations of projective materials. It is especially appropriate with data in which global, non-quantified features are utilized because the judges who do the matching are not forced to limit themselves to specified cues, but can operate with relative freedom in using various salient characteristics of the data.

However, further analysis reveals that the usual matching method (Chapman, 1934, 1935; Fiske, 1961; Secord, 1952; Vernon, 1936) demands an unnecessarily high degree of precision from the judge. He is given two sets of items, *e.g.*, handwriting samples and personality sketches of the writers, in randomized order. He tries to make the proper pairings by identifying one and only one writer for each handwriting sample. Success or failure in this attempt supposedly constitutes a test of the hypothesis that personality reveals itself in handwriting.

It is overlooked that a reduction to, say, four possible authors of each handwriting can be evaluated in such a way that it constitutes not only an adequate test of the hypothesis, but is also, as we shall point out later, a more meaningful achievement. As a rule, a judge in making his paired associations does begin by making some quick reduction to a small subset. Judges often report that the final correct matching is a matter of mere guesswork. The proper technique

should determine whether a judge's initial reductions to small subsets are more accurate than chance. Thus the judge would not have to make ultimate guesses which he may not have claimed to be able to make.

Other drawbacks of the usual matching technique stem from the statistics required by the design. Since a judge cannot use an item more than once, success or failure on any one matching influences the probability of failure on subsequent ones. The fact that one mistake necessitates another also means that the order of presentation is important: the easiest order is the one in which the simplest matchings are done first (Vernon, 1936).

All these factors create special and difficult conditions for the clinician who serves as a judge. He is forced to operate in a way which is very different from the ways he is called upon to function in practice. The requirement to make overly precise discriminations places him under stress. If he fails in the experimental task, the conclusion that what appeared as success in clinical diagnostic practice was coincidence or an illusion is unfortunately too inviting.

The present writers have designed a new form of the matching technique. Its aim is to avoid the statistical drawbacks of the one-to-one procedure and to give the clinician a chance to work in a freer manner more akin to his usual way of operating, and to avoid forcing him to do more than is claimed by the hypothesis that personality is related to, say, handwriting.

In this new design the two sets of items, personality sketches and handwriting samples, are given to the judge as before. This time, however,

¹ The authors wish to express their appreciation to Dr. Wallace Gobetz, Director, New York University Testing and Advise ment Center, whose cooperation made this study possible. They are also indebted to Irving Tumarkin for his assistance in collecting the data and preparing the personality sketches.

the judge is told that he is to associate some fixed number of personality sketches with each handwriting sample. The inclusion of the one correct match in the subset of selected sketches is counted as a success. The judge himself, after examining the data, may determine how many sketches he will use in a subset. Once he has used a sketch, he does not eliminate it for further matches; he may re-use it as many times as he likes.

Of course it is easier to get a success now than by the old matching method, and also the larger the subsets used, the more successes are apt to be obtained by chance. The authors have developed and published a test of significance for this new matching method, which takes these facts into consideration (Weinberg, Fluckiger, and Tripp, 1960).

The purpose of the present experiment is to demonstrate the outcome of a matching study under the old and the new procedures.

PROCEDURE

The data for this study were selected from the case folders of applicants to the New York University Testing and Advisement Center. Fifteen applicants were selected who seemed appropriately heterogeneous in personality. The Ss were all male and ranged in age from 25 to 40 years. As a group, they are best described as normal or mildly neurotic; the presenting problem in each case involved vocational difficulties.

First, brief personality sketches were composed from the original psychological reports. The material cited concerned the life history of the applicant and the impressions recorded by the interviewer. Next, writing samples were secured which did not describe the S or contain any other cues which might have identified him. The two sets of items were coded before they were submitted to the judges.

with graphomotor techniques served as judges. Each was given the fifteen personality sketches and the fifteen handwriting samples. He was told that for each handwriting sample he was to select those personality sketches which seemed to describe the person who had produced the sample. The number of sketches or size of the subset, was to be the same for each handwriting sample, but could be determined by the judge. The first judge proceeded to select subsets of three sketches for each handwriting, and the second judge decided to select subsets of four sketches for each. The subsets thus selected were recorded.

Then, to contrast the old procedure for matching studies with the new one, the judges were asked to go further. They now were asked to pick from each of their subsets the single personality sketch which seemed most likely to be associated with each handwriting sample. Each judge was further told not to relate the same sketch to more than one sample, so that he had to make eliminations as he went. In other words, each judge was now matching the handwriting samples and personality sketches in a one-to-one non-repetitive way, as would have been done directly by the old procedure.

RESULTS

When the first judge matched three sketches with each handwriting sample, he included the correct sketch 12 times out of 15. The second judge, who associated 4 sketches with each sample, included the correct one 11 times out of 15. In the language of the new technique, the first judge got 12 successes and the second judge, 11 successes.

To determine significance, the number of elements to be matched is designated by N , so that in this experiment $N = 15$. Also, for the new matching method, one must take into account the number of items in one group associated with each item of

Two clinical psychologists familiar

the other. The letter K will be used to designate this number. For the first judge, $K = 3$, for the second judge, $K = 4$. Table I indicates the

TABLE I—Number of Successes Needed for Significance at the .05, .01 and .001 Levels for Various Values of N and K

N	K			P
	2	3	4	
9	5	6	7	.05
	6	7	8	.01
	7	8	9	.001
11	5	6	8	.05
	6	7	8	.01
	7	9	10	.001
13	5	7	8	.05
	6	8	9	.01
	7	9	10	.001
15-22	5	7	8	.05
	6	8	9	.01
	8	9	11	.001

number of successes needed for various values of N and K at the .05, .01 and .001 levels of significance.² For the first judge 7 successes were needed for significance at the .05 level, 8 for the .01 level, and 10 for the .001 level. Since the first judge got 12 successes, his score was significant at the .001 level. The 11 successes of the second judge were also significant at the .001 level.

There are two ways of examining the results of the second stage of the experiment in which a judge had to designate the single correct sketch matching the handwriting samples. One way is merely to take into consideration the number of correct one-to-one identifications made, and to interpret the results according to a table published by Chapman (1935), for use with the old technique. The number of correct identifications made by the first judge was four (significant at the .05 level), that made by the second judge was three and was not significant.

However, the new technique allows

for analysis of the same results in a more informative way. As each judge tried to pick the single correct sketch from each of his subsets, obviously he could only be successful for subsets which included the correct sketch. The first judge had 12 chances to be correct and the second judge had 11. Since the first judge was in effect picking one sketch out of three in each set, his chance expectancy in making his final selections was four exact matches. This was what he did. Thus, although his initial reductions were significantly better than chance, his follow-up performance does not indicate that he was doing more than operating blindly. The same was true for the second judge. Given his initial reductions which were made better than chance, his subsequent number of correct matches was not significantly better than what his chance expectancy would have been. To put it another way, if after making their initial reductions, the judges had rolled dice to make their final single choice, their expectancy would have been to do as well as they did.

DISCUSSION

The results of this experiment are consistent with those of previous studies which indicate a relationship between handwriting and personality (Fluckiger, Tripp and Weinberg, 1961). The contrast between the findings for the old one-to-one method of matching and the new procedure was clear-cut. In view of this contrast, the advantages of the new method would seem demonstrated. It may be that often judges are able to make initial reductions markedly better than chance, but are not able to proceed beyond that point except by guessing. The outcome suggests that in many of the old matching studies, initial skillful decisions may have been contaminated by later random guesses.

There is one more advantage of the new technique which becomes clear

² For more extensive tables up to $N = 45$ and $K = 10$, see Weinberg, Fluckiger and Tripp (1960).

through this experiment. This is that the present type of design facilitates an investigation of the cues which judges actually use. For instance, where a judge associated four personality sketches with a handwriting sample, he necessarily excluded 11. It becomes meaningful to examine the four sketches to determine the characteristics which they have in common and which are not shared by the sketches that were excluded. An investigation of this sort cannot be done with the old method since the selection of a single case each time does not make it possible to discover commonalities.

In the present experiment, several personality sketches were matched with each handwriting sample. Therefore, the data obtained can be used to generate hypotheses as to what types of people write in a particular way. Had the judges instead been asked to match several handwriting samples with each personality sketch, the data could have been used to generate hypotheses as to what handwriting features reflect certain individual personality configurations or traits. The *E* using the new matching technique may set up his experiment in either of the two ways, depending on which best furthers his particular inquiry.

The possibility of generating hypotheses from the data does away with one of the severe limitations of matching studies. Up to now, the best a successful study of this kind could do was to support a global hypothesis of the type: Personality is expressed in handwriting in an identifiable manner. Now the cues actually used by the clinician—whether atomistic elements or complex configurations—can be specified and then subjected to a more refined experimental test. It is suggested that this feature holds promise for the whole area of projective techniques where the most difficult step is often that of making explicit whole-qualities or multiple sets of cues.

SUMMARY

The purpose of the present study was to point out some disadvantages of the usual matching method and to demonstrate the use of a new method. Fifteen handwriting samples and fifteen personality sketches were given to two judges. Each matched the two sets of data according to the new design, and each performed significantly better than chance. They then did matchings following the old method. The accuracy of both judges was markedly reduced, and one judge did not even end up matching the items significantly better than chance. The findings were interpreted as a demonstration of the superiority of the new matching method. The fact that the old method demands an unfairly high degree of precision from the judges was considered the reason for the judges' better performance by the new method. The hypothesis of a relationship between handwriting and personality was considered supported here once more. The defects of the old method were suggested as possible reasons for some research failures in the past and various reasons why the new method should be used in future matching research studies were given.

REFERENCES

- Chapman, D. W. The statistics of the method of correct matchings. *Amer. J. Psychol.*, 1934, 46, 287-298.
- Chapman, D. W. The generalized problem of correct matchings. *Ann. math. Statist.*, 1935, 6, 85-95.
- Fiske, D. W. The matching problem with multiple judges and objects. *Psychol. Bull.*, 1961, 58, 80-86.
- Fluckiger, F. A., Tripp, C. A., & Weinberg, G. H. A review of experimental research in graphology, 1933-1960. *Percept. Mot. Skills*, 1961, 12, 67-90.
- Secord, P. F. A note on the problem of homogeneity-heterogeneity in the use of the matching method in personality studies. *Psychol. Bull.*, 1952, 49, 41-42.
- Vernon, P. E. The matching method applied to investigations of personality. *Psychol. Bull.*, 1936, 33, 149-177.
- Weinberg, G. H., Fluckiger, F. A., & Tripp, C. A. A proposed variation of the matching technique. *Psychometrika*, 1960, 25, 291-295.

BRIEF REPORTS

Test - Retest Reliability of the Blacky Pictures Test

LESLIE BERGER

Franklin D. Roosevelt V. A. Hospital, Montrose, New York

LOUIS EVERSTINE

Western Psychiatric Institute, University of Pittsburgh

The Blacky Pictures (Blum, 1949) a "modified projective technique" has been developed on the basis of psychoanalytic theory to measure personality variables. In this technique, there are eleven cartoons to portray either a stage of psychosexual development or a type of object relationship within the development. Through the years much evidence on construct validity (Blum & Hunt, 1952) and on inter-scoring (Nelson, 1954) reliability has been published. However, less evidence in the way of test-retest reliability has been reported.

In the present study, the retest method of reliability will be explored.

The Blacky Pictures were given to 50 male college students in small groups. Retesting was carried out approximately four weeks later. The protocols were scored according to Blum's (1951) criteria for each of the thirteen psychosexual areas and for the four conflict patterns cutting across the various dimensions. For the former, individuals receiving an overall-dimensional score of 0 were called "low" conflict, and individuals with a score of + or ++ were considered "high" conflict.

RESULTS AND DISCUSSION

The test-retest reliability on the dichotomized individual Blacky Picture dimensions was determined by the use of the phi coefficient. Pearson r 's were computed for the Black patterns of conflict since these patterns yielded continuous scores.

An inspection of Tables I and II indicates that all 13 individual dimen-

sions and all four of the Blacky patterns of conflict yielded statistically significant coefficients. These correlations, although all are statistically significant, fluctuate considerably in magnitude. The Blacky patterns of conflict appear to yield consistently greater correlations than the individual dimensions. This might, in part, be related to the greater length of these conflict scales.

In this study we utilized the test-retest method to establish reliability measures. Since there is no alternate form to the Blacky Pictures and since the conflict scores are based on four separate sources, any split method of

TABLE I—Test-Retest Coefficients for Blacky Test Dimensional Scores and Their Significant Levels

Blacky Picture Dimension	Phi Coefficient
I Oral Eroticism	.45**
II Oral Sadism	.54**
III Anal Expulsiveness	.49**
III Anal Retentiveness	.43**
IV Oedipal Intensity	.41**
V Masturbation Guilt	.50**
VI Castration Anxiety	.47**
VII Identification Process	.54**
VIII Sibling Rivalry	.39**
IX Guilt Feelings	.47**
X Ego Ideal	.20**
XI Narcissistic Love-Object	.26**
XII Anaclitic Love-Object	.53**

* Significant at $< .02$ level of probability

**Significant at $< .01$ level of probability

TABLE II—Test-Retest Coefficients for Blacky Patterns of Conflict and Their Significant Levels

($N = 50$)

Conflict	Pearson r
Confused Sex Role	.63*
General Hostility	.58*
Disgruntled Dependency	.57*
Impulse Inhibition	.68*

*Significant at the $< .01$ level of probability

¹ The authors would like to express their appreciation to Herbert Spohn for his critical assistance.

reliability is excluded. We must, however, recognize here the limitations to the test-retest method. Familiarity with items could contribute to spuriously high correlation coefficients while changes in set and personality changes would erroneously attribute observed changes in test behavior to unreliability of the instrument. Assuming that the underlying personality which is being measured here is relatively stable, the test-retest reliability measures of this instrument, although favorably comparable with other projective instruments, are not very high. The magnitude of these correlation coefficients places certain limitations on the inferences that might be drawn from test behavior. These limitations apply especially when one is trying to infer the behavior of an individual when they are based on a single administration of the test. This is in line with Blum's (1951) original suggestion: "The objective approach to scoring is *not* recommended as the most fruitful way to interpret clinical protocols."

SUMMARY

This study was designed to explore the test-retest reliability of the Blacky Pictures scores. Blacky test records of 50 male college students who were

tested and retested within four weeks, were analyzed. All 13 of the individual Blacky dimensions, and all four of the Blacky patterns of conflict yielded statistically significant coefficients. These results are indicative of reasonably good test-retest reliability of the Blacky Pictures when administered according to standard instructions.

REFERENCES

- Blum, G. S. A study of psychoanalytic theory of sexual development. *Genet. Psychol. Monogr.*, 1949, 39, 3-102.
- Blum, G. S. *The Blacky Pictures: a technique for the exploration of personality dynamics*. New York: Psychological Corp., 1950
- Blum, G. S. Revised scoring system for research use of the Blacky Pictures. Male Form. 1951, Mimeographed copies available from G. S. Blum, University of Michigan, Ann Arbor.
- Blum, G. S. and Hunt, H. F. The validity of the Blacky Pictures. *Psychol. Bull.*, 1952, 49, 238-250.
- Blum, G. S. Reliability of the Blacky test. A reply to Charen. *J. consult. Psychol.*, 1956, 20, 406.
- Charen, S. Reliability of the Blacky test. *J. consult. Psychol.*, 1956, 20, 16.
- Lasky, J. J. and Berger, L. Blacky test scores before and after genito-urinary surgery. *J. Proj. Tech.*, 1959, 23, 57-58.
- Nelson, Sherman E. Psychosexual conflicts and defenses in visual perception. Unpublished doctoral dissertation, University of Michigan, 1954.

Received October 20, 1961

A Note on Modified Instructions for Bender-Gestalt Elaborations and Associations

MORRIS BRUCK
Pontiac State Hospital

Hutt (1945) opened up a potentially rich source of projective data from the Bender-Gestalt by introducing the idea of an elaboration and association period in addition to the standard instructions. Hutt (1945, 1953) suggested the innovation that subjects be asked to elaborate or modify the original designs in any way that rendered them more aesthetically pleasing and then give associations to the resulting production.

With our mental hospital population, it was found that these revised instructions usually failed to produce anything meaningful. Often the reaction to the revised instructions was that of anxiety, confusion and resistance. Various ways were tried to modify the revised instructions so as to bring more fruitful results. After considerable trial and error, the following instructions were found best: "Now I would like you to change the appearance of these designs in some way in order to make them into a picture of something. After you finish your drawing, write down somewhere near it what you had in mind." If patients are still uncertain about what to do, they are told: "You can add on parts, leave out parts, change them around — do anything at all to make it into a picture of something." All

of the 9 Bender-Gestalt designs are presented for elaboration.

These modified instructions have been usually found to produce meaningful drawings and associations which lend themselves to clinical interpretation and statistical compilation. The modified instructions can be readily understood and executed by patients of average intelligence from the laboring class. This was not found to be as true when the revised instructions were phrased in the manner originally suggested by Hutt. Even with persons of above average intelligence, the modified instructions introduced here seem to get more to the point and generally produce a richer source of projective material.

We also routinely ask that patients write their name somewhere on the page of the elaborations (as well as for the usual Bender performance) as an additional source of data.

REFERENCES

- Hutt, M. L. *A tentative guide for administration and interpretation of the Bender-Gestalt Test*. U.S. Army Adjutant General's School, 1945 (restricted).
Hutt, M. L. Revised Bender Visual-Motor Gestalt Test. In A. Wieders (Ed.) *Contributions toward medical psychology*. New York: Ronald, 1953, 660-687.

Received October 6, 1961

Ego Strength: A Function of the Measuring Instrument

LOREN V. COROTTO AND ROBERT H. CURNUTT
Napa State Hospital

INTRODUCTION

The term "ego strength" has been used in many different contexts, yet the term for the most part has been essentially a subjective evaluation. Recently there has been a tendency to operationally define the term "ego strength" and in a sense to delimit its meaning. This has led to several studies in which an independent behavioral criterion was used to validate the measuring instrument. Beck (1947) and Pascal and Suttell (1951) used the criterion of degree of psychological disturbance. Barron (1953) used success in psychotherapy as the criterion in the development of the ego strength (Es) scale from the M.M.P.I.

The question arises as to whether these different measures of ego strength tend to correlate with each other. An early study (Curnutt & Lewis, 1954) explored the relationship of $F + \%$ on the Rorschach with Z scores on the Bender-Gestalt (BG) and was unable to show a significant relationship between these two reputed measures of ego strength. Tamkin (1957) replicated this study and confirmed the findings of Curnutt, et al (1954). In addition, Tamkin's study examined the relationship between Barron's Es scale, Beck's $F + \%$, and the Pascal and Suttell Z score, but no significant relationship was demonstrated.

PROBLEM

Kassebaum, Couch and Slater (1959) in a study of the factorial dimensions of the M.M.P.I. explored nineteen non-clinical scales, as well as the usual thirteen clinical and validity scales, and isolated four factors. The first of these four factors, ego weakness vs.

lated in factorial studies of the M.M.P.I. It may be, then, that the previous study did not show any relationship because these extra scales had not been considered.

It is the purpose of the present study not only to attempt to correlate the BG with the Barron Es scale, but to see if there is any relationship between the BG Z score and any of the scales in the ego weakness-ego strength factor isolated by Kassebaum et al.

HYPOTHESIS

The hypothesis to be tested is that there is not a significant relationship between the BG Z score and the various scales of the M.M.P.I. which are interpreted to be an ego weakness-ego strength factor.

METHOD

Subjects: The subjects used in the present study were 105 psychiatric technician trainees, 57 males and 48 females. Each subject met the requirements specified by Pascal and Suttell (1951, pp. 100-1) in standardizing the BG Z score. The age range of the sample was from seventeen through fifty.

The ego weakness-ego strength factor isolated by Kassebaum et al. consisted of ten M.M.P.I. scales. The ten scales are as follows: Pt—psychasthenia, Sc—schizophrenia, A—Welsh's factorial scale of anxiety (Welsh and Dahlstrom 1956, pp. 264-281); Dp—Navran's dependency scale (1954); Pd—psychopathic deviate; Lp—Oettel's leadership scale (1952); To—Gough's tolerance scale (1952); Es—Barron's ego strength scale (1953); K—suppressor; and Ie—Gough's intellectual ef-

RESULTS

The results presented in Table I confirmed the hypothesis that no significant relationship exists between BG Z score and the M.M.P.I. scales isolated as an ego weakness-ego strength factor by Kassebaum et al. (1959). Since an r must be as large as .195 to be significant at the .05 level, a glance at Table I shows that not one of the ten M.M.P.I. scales examined significantly correlated with the BG Z score.

The correlation of $+.02$ between the Barron Es scale and the BG Z score confirms Tamkin's (1957) earlier study. Although the present study uses a normal population in contrast to Tamkin's psychiatric patients, the results are practically identical, i.e., r 's of $+.02$ and $+.01$ respectively.

DISCUSSION

The present results are consistent with earlier studies that were concerned with objective measures of ego strength. In short, they tend to demonstrate that three common measures of ego strength are unrelated.

These three measures, BG Z score, Barron's Es scale, and the Rorschach $F + \%$, have been operationally defined as measures of ego strength. It is argued that they all measure some aspect of ego strength. In light of the lack of significant relationships, how-

ever, one possibility may be that these instruments are measuring different, statistically independent aspects of ego strength.

Each instrument considered here may be viewed as requiring different perceptual and cognitive processes. For example, the $F + \%$ may be viewed as involving primarily visual-cognitive functioning, while the Barron Es Scale appears to involve primarily verbal-cognitive functioning, and the BG is generally considered as primarily involving visual-motor functioning.

This study fails to answer the question whether cognitive, motor, or perceptual functions are of more or less importance for processes defined as ego strength. Conceivably, this study emphasizes the probability that instruments designated as measures of ego strength do not tap *all* of these functions.

SUMMARY

The present paper explored the relationship between two operationally defined measures of ego strength, the Bender-Gestalt Z scores and ten M.M.P.I. scales isolated as an ego weakness-ego strength factor. One hundred-and-five psychiatric technicians comprised the sample. No significant relationships were found. Some implications of these findings are discussed.

REFERENCES

- Barron, F. An ego-strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, 17, 327-333.
- Curnutt, R. H. & Lewis, W. B. The relationship between Z scores on the Bender Gestalt and $F + \%$ on the Rorschach. *J. clin. Psychol.*, 1954, 10, 96-97.
- Gough, H. G. *Manual for the California Psychological Inventory*. Palo Alto, Calif.: Consulting Psych. Press, 1957.
- Gough, H. G. Special keys developed for the M.M.P.I. Unpublished manuscript, University of California, Berkeley, 1952.
- Kassebaum, G., Cough, A. S., & Slater, P. E. The factorial dimensions of the M.M.P.I. *J. consult. Psychol.*, 1959, 23, 226-236.
- Navran, L. A. A rationally derived M.M.P.I. scale to measure dependence. *J. consult. Psychol.*, 1954, 18, 192.

TABLE I—Correlations of Bender-Gestalt "Z" Scores with Selected Scales on the M.M.P.I.

Scale	Rotated factor loadings reported by Kassebaum, Gough, & Slater (1959)	
Pd	$+.01$	$+.71$
Pt	$-.03$	$+.91$
Sc	$+.01$	$+.89$
A	$-.12$	$+.88$
Dp	$-.03$	$+.85$
Es	$+.02$	$-.73$
Ie	$+.06$	$-.72$
Lp	$+.16$	$-.85$
To	$+.18$	$-.80$
K	$+.15$	$-.71$

Oettel, A. Leadership: A psychological analysis. Unpublished doctoral dissertation, Univ. of California, 1952.

Pascal, G. R. & Suttell, Barbara J. *The Bender-Gestalt Test*. New York: Grunc and Stratton, 1951.

Tamkin, A. S. An evaluation of the construct

validity of Barron's ego-strength scale. *J. clin. Psychol.*, 1957, *13*, 156-158.

Welsh, G. S. & Dahlstrom, W. G. *Basic readings on the M.M.P.I. in psychology and medicine*. Minneapolis: Univer. Minnesota Press, 1956.

Received November 2, 1961.

Relationship of Rorschach Human Percepts to Projective Descriptions with Self Reference¹

SEYMOUR FISHER

State University of New York, Upstate Medical Center in Syracuse

The significance of human content in Rorschach responses has been a matter of persistent interest. The frequency of human percepts has been studied in relationship to developmental level (Ames, *et al.*, 1952; Ames, *et al.*, 1959); cultural norms (Klopfer, 1956); self concept (La Fon, 1954); personality pathology (Rickers-Ovsiankina, 1954; Holtzman, *et al.*, 1961) and numerous other variables (Klopfer, 1956). Many of these studies have suggested that the human response is an indicator of the degree to which an individual has been able to establish a meaningful and acceptable sense of identity. It may be conjectured that the projection of human concepts on to an ink blot reflects a positive identification with the role of person or social entity. The following specific formulations derived from this general view were investigated in the present study:

1. The greater the frequency of human responses the more positively and acceptingly does the individual view himself as a person.
2. The occurrence of human responses is negatively related to confusion or conflict about sexual identity.
3. Number of human responses is negatively linked with concern about personal vulnerability or fragility.
4. Number of human responses is negatively related to a sense of being childlike or immature.

PROCEDURE

The human response scores were tabulated from Rorschach protocols which had been secured on an individual basis. One departure from stan-

dard procedure was to request subjects to give three responses to each blot. Thus a total of 30 responses was uniformly obtained. This was done in order to control for response total. Each clear reference to a human percept was designated as H; and the H score was equal to the sum of such references.

In order to measure the role variables mentioned in the hypothesis a technique was developed based on the individual's projective descriptions of his own image in a mirror while wearing various types of masks. It was operationally presumed that if an individual is requested to describe his own distorted image in an unstructured, somewhat stressful situation he will reveal significant aspects of his attitudes toward himself. The actual procedure used involved bringing the subject into a totally dark room and placing him four feet from a full-length mirror. He was told that a series of masks would be placed on his face and that he would briefly see himself in a mirror while wearing each mask. It was requested that following his view of each mask he describe his appearance in complete detail. Such descriptions were written down as verbatim as possible. During the actual course of the procedure the subject was told that when each mask was placed on his face he was to keep his eyes closed until a preparatory signal was given. He was then to open his eyes; look straight in front of him; and prepare to see the image of his face which would be briefly illuminated in the mirror. When his face was no longer illuminated, he was to close his eyes and describe what he had seen. The mirror at which the subject looked was lighted by two

¹This study was partially supported by USPH grant M1578.

lights (75 watts) for one second by a timer which the experimenter started one second after signaling him to open his eyes. The masks used consisted of four male and four female faces that were presented in random order. They were rubber masks quite life-like in appearance, but yet with sufficiently vague expressions that they were likely to lend themselves to projective elaboration.

The descriptions obtained from each subject were scored for four variables:

1. Degree of self satisfaction was equated with the number of masks referred to in a positive or satisfied manner. If one or more terms with such positive connotations were used in relation to any given mask, it was scored in the positive category. Terms like "nice," "happy," "friendly," "pleasant," "successful," "good humored," "strong," and "good thinker" exemplify what were considered to be positive references. In other words, a description was labeled as positive when in ordinary usage it would be regarded as having complimentary or approving significance. Interscorer agreement between two judges for 30 protocols was 89 percent.

2. An index of confusion about sex role was derived in terms of the number of instances in which the sex of a mask was misidentified.

3. Sensations of personal vulner-

ability were evaluated by means of a count of the number of masks which were described as portraying injury or disfigurement (e.g., cut, scarred, dead, crushed, broken).

4. A score designed to reflect feelings of immaturity and childishness was derived from a count of the number of masks which were perceived as not being of adult age (e.g., child, baby).

SUBJECTS

There were 63 subjects of whom 33 were female and 30 male. Their median age was 20 and median educational level 14 years. They were recruited from sororities and fraternities; and were paid a fee for their participation.

RESULTS

The median number of H responses was 6, with a range of 1-15. With regard to the mask indices, the median Positive Attitude score was 0 (range 0-5); the median Sex Misidentification score 1 (range 0-4); the median for Vulnerability 0 (range 0-4); and the median for Immaturity 0 (range 0-2). The distributions of scores were generally very skewed and so non-parametric data analysis (viz., chi square) has been employed throughout. There were no sex differences except for the Sex Misidentification score. Women obtained significantly lower scores (.001 level) for this variable than did men.

TABLE I—Chi-Square Tests of Relationships Between Rorschach Human Content Scores (H) and Mask Description Indices.

Variables	Rorschach	H Responses		X ²	Significance Level
		H ^a	L		
Positive Attitude	H	28	17	6.1	.02-.01
	L	5	13		
Sex Misidentification	H	15	21	3.9	.05
	L	18	19		
Vulnerability	H	6	13	4.7	.05-.02
	L	27	17		
Immaturity	H				NS ^b
	L				

^aH = Above median

L = At median or below

^bNS = does not attain .05

Table I indicates that three of the four results were in the predicted direction. Subjects with high H scores on the Rorschach were significantly higher than subjects with low H scores on the Positive Attitude ($X^2 = 6.1$; $p = .01$) variable and significantly lower on the Sex Misidentification ($X^2 = 3.9$; $p = .05$) and Vulnerability ($X^2 = 4.7$; $p = .05-.02$) indices. The Immaturity score did not differentiate between subjects in the high and low H groups. It should be noted that none of the results quite attained significance within the individual sex groups.

DISCUSSION AND CONCLUSIONS

The results support the view that the ability of an individual to project human percepts on to the Rorschach blots is related to the adequacy of his role identity. Within the context of the present study one might say that high H goes along with a positive orientation toward self, a consistent perception of one's sex identity, and minimal feelings of vulnerability or fragility. Of course, no data have been obtained which demonstrate that the mask descriptions do, indeed, convey information about a subject's role or identity. Strictly speaking, one is only in a position to say that H responses bear certain significant relationships to the way in which subjects describe masks worn on their own faces in a situation which by its unstructured character encourages projection. But even if one limits the interpretation of the mask descriptions to this opera-

tional level, one is still left with an interesting and suggestive series of relationships.

SUMMARY

This study explored the general hypothesis that the ability to project human responses on to the Rorschach blots represents a positive identification with the role of person or social entity. Sixty-three subjects were individually tested with the Rorschach test. The frequency of human responses in their protocols was related to four indices of presumed self concept derived from their self descriptions of their own mirror images while wearing a series of masks. In general, the results were supportive of the hypothesis.

REFERENCES

- Ames, Louise B., Learned, Janet, Metraux, Ruth W. & Walker, R. N. *Child Rorschach responses*. New York: Paul B. Hoeber, Inc., 1952.
- Ames, Louise B., Metraux, Ruth W., & Walker, R.N. *Adolescent Rorschach responses*. New York: Paul B. Hoeber, Inc. 1959.
- Holtzman, W. H., Thorpe, J. S., Swartz, J. D., & Herron, E. W. *Inkblot perception and personality*. Austin, Texas: University of Texas Press, 1961.
- Klopfer, B. *Developments in the Rorschach technique*, Vol. 2. New York: World Book Company, 1956.
- La Fon, F. E. Behavior on the Rorschach test as a measure of self acceptance. *Psychol. Monogr.* 1954, 68 (No. 10), Whole No. 381, 1-14.
- Rickers-Ovsiankina, Maria A. Longitudinal approach to schizophrenia through the Rorschach method. *J. clin. exp. Psychopath.*, 1954, 15, 107-118.

Received October 9, 1961

A Simplified Method of Scoring Rorschach Content for Dependency

EUGENE E. LEVITT AND BERNARD LUBIN

Indiana University Medical Center

and

MARVIN ZUCKERMAN

Brooklyn College

The importance of the construct, dependency, to personality theory as well as to applied areas like psychotherapy is acknowledged by most workers in these areas. It ranks along with anxiety and hostility as broad constructs which are widely regarded as basic to clinical research and practice. The development of measures of dependency, however, lags well behind that of anxiety and hostility.

The present paper is intended as a contribution to the area of objective measurement of dependency. It presents a relatively simple method for scoring Rorschach content for dependency. Much further research with this system is needed, and it is hoped that its publication will stimulate such endeavors.

The scoring system is essentially a simplification of that of De Vos (1952, 1955). The revised, elaborated form of his system admittedly includes percepts within dependency which could also be regarded as indicators of either anxiety or hostility (De Vos, 1955). Such concepts have been removed in our simplified system. De Vos' nine primary categories and 64 subcategories have been reduced or condensed into three primary categories and 16 subcategories. Basically, however, the inferential logic used by De Vos to develop his system is retained in ours.

THE SCORING SYSTEM

The method of scoring Rorschach content for dependency is reproduced in its entirety below.

General instructions: Count 1 for each response indicating dependency. For Categories I and II, base scoring on the *idea*, not on specific language.

Category III scorings are based on specific wording.

I. Movement Responses

A. *Animals or humans in movement expressing submissiveness, passivity, dependency, helplessness.*

e.g.: Bowing, kneeling, leaning, reclining, resting, clinging, hanging onto something, sleeping, being helped up or down, praying, floating, being picked up, drooping, dangling

B. *Oral Activities*

e.g.: Smoking, blowing, drinking, eating, nursing, sucking, being fed, holding something in the mouth

C. *Childish activities*

e.g.: Playing patty-cake, hop-scotch; animals (but not humans) dancing

II. Content

A. *Diminutive animal figures or animal toys*

e.g.: Calf, lamb, colt, cub, puppy, teddy-bear, toy dog

B. *Reference to oral anatomy, animal or human*

e.g.: Mouth, lips, gums, tongue (do not score teeth)

C. *Oral suppliers*

e.g.: Breasts, nipples, udders, baby bottle, preparation of food for eating

D. *Food*

e.g.: Ice cream, carrots, dessert, cooked chicken or meat (do not score raw meat)

E. *Human content of the following types*

1. Diminutives: Baby, child, infant

2. Human-like diminutives: gnome, dolls, gremlin, fairy, imp, pixie, goblin, puppet, marionette (do not score witches)

3. Human-like figures originating in, or popularized by, the literature of childhood: Santa Claus, Wizard of Oz, Bugs Bunny, Donald Duck, certain comic strip characters

1. Religious figures: Moses, Christ, priest, monk, nun
 5. Figures tied together: Siamese twins, two dogs with feet tied together
- F. *Contrasts in size of figures or an emphasis on size from the point of view of the observer*
 e.g.: A giant, like you are looking up at him; a big insect about to pounce on a little insect.
- G. *Religious objects or buildings of Western religions.*
 e.g.: Church, crucifix, altar
- H. *Authority symbols*
 e.g.: Emblem of the United States, officer's insignia, crown, coat-of-arms
- I. *Toys (including Halloween mask)*
- III. *Style*
 Any verbalization using childish-like language, e.g., scary, wee-bitty, etc.

RELIABILITY OF SCORING

A group Rorschach was administered to the sophomore class of student nurses at the Indiana University School of Nursing ($N = 72$) in 1959, and again to the sophomore class of 1960 ($N = 74$). The cards were projected on the screen using an opaque projector, and subjects were permitted three minutes per card in which to write their responses.

Each protocol in both samples was scored by the same two raters, neither of whom had, or was working toward, a graduate degree in psychology.¹ The mean number of dependency responses by each rater for each group, the mean rating for each group, and the correlation between ratings, are shown in Table I.

TABLE I—Reliability of Dependency Content Scoring

	Mean Number of Dependency Responses by Rater:			
	A	B	Mean Inter-rater r	
1959	5.86	6.13	5.99	.83
1960	4.87	9.22	7.05	.76

1959 5.86 6.13 5.99 .83
 1960 4.87 9.22 7.05 .76

The inter-rater correlations of .83 and .76 compare favorably with the

average reliability estimate of .68 reported by De Vos (1952) using expert judges. The improved reliability probably reflects the greater ease of scoring of our system as compared to De Vos' original formulation.

EVIDENCE FOR VALIDITY

De Vos' original dependency scoring system did not differentiate among normals, neurotics, and schizophrenics (1952), but it is questionable whether it should have been expected to differentiate these groups. It is certainly questionable whether our revised system, with overlapping anxiety and hostility percepts removed, should be expected to differentiate emotional disturbance from normality. As a first field test of our system, we chose a pair of contrasted groups for which the logic of a difference in dependency appears more tenable, namely volunteers and nonvolunteers for an hypnosis experiment.

The logic of the choice stands on fairly sound theoretical grounds. The involvement of the personality factor, suggestibility in various aspects of hypnotic behavior is well established. Weitzenhoffer (1953, p. 71) suggests further that "most people who volunteer as subjects do so because they are highly suggestible in the first place and the request of the hypnotist for volunteers acts selectively as a suggestion upon these very individuals." Recent evidence by Zuckerman and Grosz (1958) indicates "that a person who is suggestible is likely to be a person with strong dependency needs . . ."

The hypothesis that volunteers for an hypnosis experiment should manifest a greater degree of dependency than nonvolunteers appears to be reasonably tenable. Clinical observation suggests further that the dependency need itself should exist on a deep-lying level of the personality so that the subject would be unaware of his true motivation in volunteering. This is the level, of course, which is presumed to be tapped by a projective technique.

¹ The authors are indebted to Harry Brittain and George Petoe for rating the protocols.

The contrasted groups in our validity study were taken from the sophomore class of student nurses at the Indiana University Medical Center in 1959 and again in 1960. The subjects had been asked to volunteer privately for an experiment in which they would be hypnotized. The group Rorschach upon which the data of Table I are based had previously been administered to all students.

The mean percents of dependency responses are shown in Table II.

TABLE II—Rorschach Dependency Content Scores of (% of R) of Volunteers and Nonvolunteers for Hypnosis Experiments

	N	Mean%*	SD
1959: Volunteers	30	25.10	10.5
Nonvolunteers	42	18.24	6.9
1960: Volunteers	38	25.99	9.7
Nonvolunteers	36	21.31	9.1

*1959 $t = 3.50$, $P < .01$

1960 $t = 2.17$, $P = .05$

The volunteers have the higher mean dependency percentage scores in both years. The difference for the 1959 contrast yields a t of 3.50, $P < .01$. The t for the difference in the 1960 groups is 2.17 which is significant at the 5% level. The results are therefore in accord with the hypothesis that volunteers for an hypnosis experiment should have more deep-lying dependency needs than nonvolunteers, and therefore provide some evidence of construct validity for our scoring system.

ADDITIONAL DATA ON THE SCORING SYSTEM

There is also available certain additional information concerning dependency scores furnished by our system, though they do not necessarily cast light on either the validity or the reliability of the system. These findings are summarized from the article by Zuckerman, Levitt, and Lubin (1961), and are based on the 1959 sample only.

The correlation between the de-

which is not significant, suggesting that our dependency score is not related to intelligence. The dependency score did not load on any of the four factors which appear to be involved with dependence-independence as a personality characteristic. (These factors were made up almost entirely of scores from verbal inventories). A subsample of 20 of the Rorschach protocols were divided into high and low dependency groups on the basis of the dependency score. Two expert judges, working independently, agreed with the classification for 80% of the cases, using global judgments of dependency based on the entire protocol. The probability of this occurring by chance is less than .02.

SUMMARY

A simplified system of scoring Rorschach content for dependency, based on that of De Vos, is presented. The reliability of the scoring, even by non-expert judges, seems adequate. Data from two replications of the study in which it was shown that dependency scores for volunteers for an hypnosis experiment were significantly greater than for nonvolunteers, are presented as some evidence for the construct validity of the scoring system. It is hoped that the greater simplicity of our system as compared with that of De Vos will stimulate experimental study of the dependency factor using Rorschach content as a measure.

REFERENCES

- De Vos, G. A quantitative approach to affective symbolism in the Rorschach. *J. proj. Tech.*, 1952, 16, 133-150.
- De Vos, G. *Manual of criteria for scoring affective inferences*, 1955. (unpublished)
- Weitzenhoffer, A. M. *Hypnotism: an objective study in suggestibility*. New York: J. Wiley, 1953.
- Zuckerman, M. and Grosz, H. J. Suggestibility and dependency. *J. consult. Psychol.*, 1958, 22, 328.
- Zuckerman, M., Levitt, E. E., and Lubin, B. Concurrent and construct validity of direct and indirect measures of dependency. *J. consult. Psychol.*, 1961, 25, 316-323.

A Simple Method for Measuring the Area of a Figure Drawing

JEROME D. PAUKER
State Hospital No. 1, Fulton, Missouri

When reference is made to the measured size of a figure drawing, it is usually the height of the drawing, head-to-toe, which has been measured (e.g., Hammer, 1958). True size, however, should refer to the *area* which a drawing covers, and area will be, to a certain extent, independent of height. An easily obtained, exact measurement of area can be made with an instrument called a compensating polar planimeter, but this device is relatively expensive. This paper describes a simple, inexpensive, easily-made plastic grid which can be used to get a measure of area that correlates very highly with true area, and compares this method with measurement of height as an estimate of true size.

MEASURING GRID

The measuring grid consists of a piece of clear plastic, 9 by 11½ inches. A quarter-inch border is drawn to provide an 8½ by 11 inch frame. Vertical and horizontal lines are drawn within the frame, each line one-quarter inch away from its parallel neighbors, to provide a grid of quarter-inch squares. The lines can be scored with a metal point and ink can be dropped into the grooves to provide more permanent marks.

METHOD OF MEASUREMENT

The measuring grid is placed over the drawing so that the 8½ by 11 inch page is lined up with the grid frame. Then the number of squares which are entered by any part of the drawing are counted. Any square which is even slightly entered by a line, or which covers an area enclosed by the boundary lines of the figure, is counted. This count is the measure of area.

CORRELATION OF GRID AND OF HEIGHT MEASURES WITH PLANIMETER MEASURES

Sixty male college students¹ were asked to draw a picture of a person on a sheet of 8½ by 11 inch paper. The areas of these drawings were measured with a K&E compensating polar planimeter (model No. 4236M). The polar planimeter measures a plane area in any form. A tracer point is run around the edge of the figure and the distance which a measuring wheel has revolved during this process is read off. The model used here reads directly in square centimeters to a tenth of a square centimeter.

The areas of the drawings were also determined with the measuring grid in terms of quarter-inch squares, and the heights were measured to the nearest eighth of an inch.

The mean measurements by planimeter, grid, and ruler are, respectively, 42.57 square centimeters, 141.40 quarter-inch squares, and 5.34 inches, with standard deviations of 34.16, 97.25, and 2.00, respectively.

The grid and planimeter measures correlate .995 (Pearson *r*). The height and planimeter measures correlate .875.

From these correlations it can be seen that the height of drawings does, indeed, correlate very well with area, but that when a measure of area is required that will correlate almost perfectly with actual size, then the use of the described measuring grid will provide a simple and inexpensive way of obtaining it.

¹ The subjects were undergraduates in General Psychology and Sociology courses at the Normandy Residence Center of the University of Missouri. The cooperation of Mr. C. E. Potter, Mrs. I. B. Mathes, and Dr. D. G. McDonald is acknowledged with appreciation.

SUMMARY

A grid of quarter-inch squares drawn on clear plastic can be used to obtain a measure of the area of figure drawings by counting the number of squares which the drawing enters. This measure correlates .995 (Pearson r) with area measured with a compensating polar planimeter. The head-to-

toe length of the figure also correlates respectably with the planimeter measure ($r = .875$).

REFERENCE

Hammer, Emanuel F. *The clinical application of projective drawings*. Springfield: Charles C. Thomas, 1958.

Received August 14, 1961

The Utilization of the Patient-Examiner Relationship in Intelligence Testing of Children¹

MILTON F. SHORE

Worcester Youth Guidance Center, Worcester, Massachusetts

It is generally accepted in our study of the growth of the individual that he goes through many stages of development. In a similar fashion, our understanding and views regarding the nature of intellectual functioning from performance on intelligence tests have also gone through a number of stages. An analysis of the historical development of our ideas about intelligence reveals three such stages:

1. The stage of classification, quantification, and measurement. This stage was characterized by pure empiricism.

2. The stage of the clarification and understanding of the theoretical, projective, and inter-personal aspects of the intelligence test situation.

3. The current and still developing stage—the active utilization of the inter-personal and projective material within the context of the intelligence test situation in order to explore certain aspects of intellectual and personality functioning and their inter-relationship.

In this paper I shall confine myself to the Wechsler intelligence scales, but the statements I make are equally relevant to other intelligence tests. The Wechsler, however, is unique in that it has elicited such wide interest and, as a result, has led to a vast amount of research and theoretical work. Although many of my references will be from the work on the adult scales, I shall illustrate with examples from the WISC.

The initial stage in the development of tests of intelligence, arising from the early interest in faculties and

their measurement, was the construction of a careful statistical model for valid and reliable measurement of the variable. This model was based on the originator's theoretical conception of the nature of intelligence. The goal of Wechsler, for example, was to obtain a valid and reliable measure of a person's achievement on a variety of tasks, verbal and performance, as compared with other people of his chronological age.

Personality factors during this first phase were considered obstructions in a manner similar to the concept of "stimulus error" in early studies using introspection as a method. The obstructions, it was believed, interfered with the true expression of a person's abilities by affecting his performance in many ways. An example of this approach was Wechsler's listing of non-intellective factors affecting performance, (e.g., special skills, drives, incentives, etc.) without specifying the nature of their influence on the test performance.

In line with this concern with achievement scores came many attempts to isolate various patterns of scores and relate these patterns to particular deficits within certain diagnostic groups. This approach, especially in its early phases, was noticeably lacking in theory, with much of the work being purely empirical in character. The approach began with the Deterioration Index in the adult scales (based on those subtest scores which drop with increasing age), and led to the highly complex scatter analyses of Rapaport.

The lack of success of attempts to validate scatter analysis, especially when applied to the individual patient, has been noted in all summaries

¹ Presented at a Symposium entitled "Current Theory and Practice with the WISC" at the Annual Meeting, American Psychological Association, New York, 1961.

of research on the Wechsler intelligence scales. Littell (1960), for example, in his recent article on research with the WISC has concluded that the scatter analysis approach to the WISC has been found to be of limited value. The reasons for this lie in the fact that the approach is weak on both theoretical and statistical grounds.

The second stage in the growth of our understanding of intellectual processes as measured by intelligence tests can be said to have begun with Rapaport's formulation (1946) of the "projective hypothesis" which stated that "every reaction of a subject is a reflection, or projection of his private world." This included intelligence tests, which are highly structured, as well as the projective tests whose main purpose is to elicit the unique ways a person deals with the world.

Four important trends resulted from Rapaport's formulation of the "projective hypothesis":

1. An attempt was made to develop a theoretical rationale for each of the subtests. This permitted more meaningful and significant interpretations of the person's performance. However, there remain many problems within this area.

Rapaport's rationales, which were the primary impetus to this approach, were based on intuitive appraisals. Although they have been accepted by many as fact, his rationales remain only hypotheses with much work still to be done to test their validity.

With regard to subtest rationale, the WISC also presents a unique problem which is not present in the adult scales. In working with children we are forced to consider the effect of development on the various psychological functions measured by the test. For example, does a sub-test mean the same thing and measure the same functions at different age levels? Unless such a question can be answered affirmatively, the test must be used before

applying Rapaport's rationales for the adult scale to the WISC.

2. The focus on process as well as achievement. Thus, interest arose not only in whether the answer to an item was correct or not, but also in how the person arrived at his answer. This dynamic view of intellectual functioning placed intelligence within the context of total personality functioning and led to the description of the projective aspects of the intelligence test performance.

3. Many attempts were made to organize and make theoretically consistent the style of an individual's performance on intelligence tests. This has taken many forms depending upon the particular theoretical orientation of the writer. For example, one approach was trait classification based on the particular requirements of the task and how the individual differed from others in his approach to the task. More complex ways of looking at performance on intelligence tests have been based on psychoanalytic theory and involve the evaluation of defensive maneuvering, interference of regressive material, perception of object relations, etc.

4. Recently there have been many suggestions that the interpersonal aspects of the intelligence test situation be considered. Schafer (1954) has described the complex aspects of the interpersonal situation in testing with specific reference to the Rorschach. Many of his comments are relevant to the more highly structured intelligence test situation. For example, how does the patient perceive the examiner? — as a friend? as a person prying into his secrets? as a doctor out to help him with his difficulties? or as a teacher testing him to get a grade for promotion? How does the examiner's attitudes towards the patient, towards the test, and towards himself influence the patient's performance? For example, does the examiner wish to get his job over as quickly as possible?

Does he feel he has something worthwhile to offer the patient?

This area of the interpersonal aspects of the testing situation is especially important in work with children for two reasons:

First, because they have not as yet developed an adequate defensive structure, children are particularly sensitive to the attitudes of others especially if these people are adults who are asking them to perform in a particular manner. Second, children bring with them a variety of experiences with adults which they have as yet not adequately integrated. Therefore, unlike adults, they are frequently unable to isolate the testing situation from their other experiences with adults.

The third stage in our understanding of the intellectual functioning is currently in the process of development. It is where the examiner actively explores within the intelligence test situation certain aspects of intellectual functioning, of the interpersonal situation, and/or various areas of personality structure. Within this context the intelligence test is seen as a semi-structured interview situation requiring the use of certain ego functions by the patient, but offering a wide field for exploration in order to obtain valuable diagnostic material.

This view of testing was first presented at the 1958 Annual meeting of the American Psychological Association by Leventhal, Rosenblatt, Gluck, and Slepian (1962). It was their belief that, although the relationship between patient and examiner had now been recognized, the relationship could be actively used: (1) to explore certain areas for diagnostic purposes, (2) to utilize the situation to establish a clinical contact where the examiner is seen as a person who wants to understand and help, (3) to help allay a great deal of the patient's initial anxiety so that he is able to function closer to his optimal level (the tests, of course, are scored and the quanti-

tative results used in evaluating the patient's total performance).

What is the relevance of this approach to intelligence testing with children?

1. The pre-test preparation. Not only are questions asked like How does the child perceive the testing situation? but active efforts are made to clarify and correct any unusual views the child may have. There is active exploration of the purpose of the testing and the child's previous experiences with testing. The current test situation is separated from tests given in courses, needles given in doctor's offices, etc. It may be necessary to reassure the child about the confidentiality of the relationship. Usually it is also made clear to the child at the beginning what is to be done with the tests after they are completed, and how he may find out about the results. In this way an effort is made to handle some of the initial anxieties about testing. We have found that such an approach has served to relax many children so that they are able to respond to the tests at an optimal level, an aim we set ourselves frequently when testing disturbed children.

2. Handling of feelings which arise during the administration of the intelligence test itself. The kinds of feelings arising during the testing, of course, cannot be predicted. How these are handled will also depend upon the examiner's clinical understanding of the case. Thus, if a child cries when attempting an item, the examiner must judge whether this is true helplessness or whether it is a histrionic attempt to manipulate the situation. The examiner's response is quite different in these two situations. Also when a child gives up the examiner must decide whether or not he will accept it or whether he will attempt to get the patient to respond.

3. The giving of help. The psychologist does not hesitate to give help at times during testing. In addition to the usual support and encouragement

he may even offer suggestions or participate in the completion of the task.

(Where the examiner has participated actively the item is not scored correct. Where some help has resulted in a correct solution, this is noted.) How the child responds to help may provide valuable information about the way he works on intellectual tasks. For example, is the child able to profit from the help given on one item in doing another item, or does he proceed in a stereotyped, rigid manner?

4. The exploration of various intellectual functions. Although the directions for administering the tests are adhered to as much as possible, they do not bind the examiner. We should remember that the WISC was not standardized on children with emotional problems which constitute the majority of children we see for testing. We, therefore, employ some of the techniques of Piaget and Edith Meyer Taylor (1959) who do not hesitate to deviate from standard instructions when it is necessary to get a clear understanding of how the child functions in various intellectual tasks.

5. The exploration of various aspects of personality structure. Idiosyncratic breakthroughs in response to a question are not accepted without exploration. For example, if there is a response which indicates that reality testing may be impaired, the child may be asked if he really believes things can be that way, in order to see if he is able to check his performance with some judgment. In a similar manner, we explore certain defensive maneuvers with regard to the test items and may even discuss the relationship between the patient's test performance and his behaviour in other situations such as in school. We may point out, for example, that the child has done well on the tests so why is he having trouble in school?

6. Exploration of interpersonal areas of functioning. Certain patterns of behavior toward the examiner in the completion of certain tasks may

be taken up. For example, excessive need for reassurance and support, sensitivity to criticism, playing dumb, etc.

7. Exploration and clarification of certain areas of fantasy. The psychologist may take up with the child certain dynamic factors that present themselves in the intelligence^W test responses. For example, he may note that the child seems pre-occupied with anger, or that he seems to constantly feel that punishment is near.

In no way do I want to imply that in each case we cover all the aspects listed. How far one goes in exploration depends on the personality of the particular examiner and his understanding of the situation. There may be times when a patient is so extremely loose and tangential that exploration would be harmful. In these cases the test should be given in a highly structured fashion.

In other words, the approach suggested above utilizes the clinical insights learned through extensive clinical training in psychotherapy and diagnostic work. The Wechsler is seen in this context as a clinical and diagnostic tool rather than only as a psychometric instrument or as a test with projective aspects.

In summary, the use of the Wechsler intelligence scales has gone through three stages of growth which can be labelled as psychometric, projective, and the latest, clinical. In no way do I want to imply that one stage is less important than another. Repeating our analogy to the individual and his development, each stage depends on the preceding one and integrates the previous stage within the total context of the new one. In fact, it is my belief that it is precisely because the Wechsler has gone through these stages that it continues to have such great value in our continued efforts to understand the nature of man.

REFERENCES

- Leventhal, T., Gluck, M. R., Slepian, H. J., & Rosenblatt, B., The utilization of the psychologist-patient relationship in diag-

- nostic testing *J. Proj. Tech.*, 1962, 26, 66-79.
- Lattell, W. M., The Wechsler Intelligence Scale for Children: A review of a decade of research. *Psychol. Bull.*, 1960, 57, 132-157.
- Rapaport, D., *Diagnostic psychological testing*. Vol. 1. Chicago: Yearbook Publishers, 1946.
- Schaffer, R., *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- Taylor, Edith M., *Psychological appraisal of children with cerebral defects*. Cambridge: Harvard University Press, 1959.

Received October 11, 1961

BOOK REVIEWS

Buss, A. H. *Psychology of Aggression*. New York: John Wiley and Sons, Inc. 1961, Pp. X + 307, \$7.95.

The control of aggression, both individual and collective, is increasingly becoming one of the most important areas of concern for clinical psychologists. This scholarly volume by Arnold Buss, a clinical psychologist at the University of Pittsburgh, is an outstanding contribution to the synthesis of whatever knowledge we now have about the psychological mainstays of this behavior. The book goes far beyond both the slim volume of Scott (Scott, J. P., *Aggression*, Chicago: University of Chicago Press, 1958) and the McNeill monograph (McNeill, E. B., *Psychology and aggression*. *J. conflict res.*, 1959, 3, 195-293.) in integrating empirical and experimental findings with theoretical contributions. Buss leaves no bit of evidence unturned. One of the outstanding contributions is his citation of many hitherto unpublished findings, including a plethora of Ph.D. theses which otherwise would never have gotten out of the stacks. All of these studies he submits to critical examination from the viewpoint of an unremitting reinforcement theorist, rarely avoiding negative evidence however, and always attempting to conciliate and integrate discrepant findings.

The book is divided into three parts, the second of which, covering theories of aggression, methods of measuring it, and psychopathological manifestations, will probably be of particular interest to readers of this journal. However, they would do well to read the other sections carefully. Part I deals with laboratory studies of aggression; and what evolves actually is a "how to do it" manual, the advantages and disadvantages, the problems, pitfalls and payoff of manipulative studies. Special attention is given to the problems of measuring aggression in the laboratory and recommended procedures for overcoming the different sets with which the individual comes there. This should be helpful to anyone doing research in this area.

Also of service is the distinction Buss draws between aggression, "an instrumental response that administers punishment"; anger, "an emotional reaction with prominent autonomic and skeletal-facial components"; and hostility, "a negative attitude which is

terms of implicit verbal responses". Any one of these three elements can occur in the absence or presence of the others. Such distinction should help in understanding seemingly contradictory findings. Although Buss makes quite a point of eliminating the notion of intent from his definition of aggressive behavior, it seems to us that when he eliminates both accidents and acts with a socially desirable effect such as a dentist drilling a tooth, a parent punishing a child or a doctor administering a shot, he is back in the same operational dilemma since judgments still have to be made by someone (hopefully more than one) as to whether a given act is accidental or when it is socially desirable. Is punching a bully in the nose an aggressive act or does it have a socially desirable effect? Is a hunting incident, wherein the sportsman aims for a deer and fells his companion, an aggressive act or is it an accident? There might, after all, be no reinforcement history for these specific responses to which we can turn in making what is essentially a value judgment which Buss decries. And when he includes any "attempt to deliver noxious stimuli" (p. 5) in his definition, he sinks deeper into a morass of judgments. When is the "playful" feint, which may or may not connect, an aggressive act or when is it all in fun? It seems that Buss has unnecessarily gotten away from a strictly behavioral definition and has not eliminated intent or the need for value judgments at all.

In his treatment of punishment, Buss makes a literal translation from reinforcement theory as developed in the rat laboratory to human behavior. He predicts that punishment of aggression will lead to its inhibition because of the anxiety which is aroused by the prospect of punishment in subsequent similar situations. In carefully controlled laboratory studies with college student subjects this often happens. But such a formulation disregards the monotonous finding of field studies to the effect that increased punishment for aggressive behavior by socializing agents is related to increased aggression on the part of the subject. The results of such studies of parent-child relations suggest that punishment for aggression acts more as an instigator than a negative reinforcer, or else provides the child with an aggressive model whose behavior can

Part III deals with the development of aggression in children, about which there is still little good information other than in regard to age and sex differences, despite the number of studies and techniques which abound; and with the relationship between aggression and prejudice. In a keen analysis of many conflicting findings the author thoroughly demolishes the scapegoat theory of prejudice except in the case of the extreme bigot, and offers as an alternative approach one of social learning, or imitation and reinforcement (albeit with no reference to Rodgers and Hammerstein).

As noted above, it is with Part II, especially the chapter on Projective Tests, that readers of this journal will be most concerned. In his discussion, Buss limits himself to the Rorschach and TAT and, while his treatment of the research literature concerned with these two methods is exhaustive, there are other techniques which might have also been profitably covered, such as the Picture Frustration Study which is directly concerned with aggressive responses. Buss' assessment of the value of the Rorschach as a measure of aggression is not optimistic. Scoring formal categories has met with no success in predicting to an outside criterion; scoring of hostile content has been a little more successful, relating to long time aggressive trends in the individual, although unaffected by laboratory manipulation of aggressive need. The Thematic Apperception Test, however, has been the instrument without peer in reflecting such transient changes in aggressive need. In summing up the results, pro and con, of an old controversy, Buss concludes that the overwhelming weight of evidence points to the TAT as directly representational of overt aggressive behavior and not as compensational for it; the drainage hypothesis has not been substantiated. In general, Buss' conclusion from a review of research on the utility of projective techniques in the assessment of aggression is that they "reveal only the aggressive trends that subjects, if given the opportunity, can verbalize" (p. 155). They do not yield "underlying deep aggressive trends"; at least with normal subjects, as much information can be obtained from a questionnaire. However, since none of the studies quoted utilized adequate outside criteria of such unconscious trends, nor did they even try to predict to them, it is hard to see how so sweeping a conclusion is completely justified now. Scoring procedures based on clinical insights may yet be developed which will tap such latent aggressive content and hopefully criteria will also be

developed against which these scoring systems can be adequately validated.

One oversight of Buss has to be mentioned. He cites the often quoted, but unfortunately unpublished, study by Pittluck (Pittluck, Patricia. The relation between aggressive fantasy and overt behavior. Unpublished Ph.D. thesis, Yale University, 1950), and makes a number of criticisms, some justified; however, he fails to mention that she was the first TAT researcher to suggest that in order to sharpen prediction from TAT stories to overt behavior it is necessary to take into account not only indications of aggressive need, but also manifestations of anxiety inhibiting its expression, a point he makes much of in other TAT studies which are cited, as well as in different contexts.

Despite the minor criticisms noted here, this reviewer feels that Buss has made an outstanding contribution in critically summarizing all available research evidence as regards psychological variables in the development of aggression, in pointing out the gaps in our knowledge and in indicating promising leads for future research. His evaluation of the present status of projective techniques in the assessment of aggression is fair. This book is highly recommended to all clinical psychologists interested in problems of aggression either in their practice or in research.

LEONARD D. ERON
Rip Van Winkle Foundation
Hudson, New York.

✓ **Eiduson, Bernice T.** *Scientists—their psychological world.* New York: Basic Books, Inc., 1962. pp. 299, \$6.50.

This book reports some of the data and conclusions from a clinical psychological study of 40 male physical scientists who are actively engaged in scientific research. It is written in an informal, personal style and aimed at scientists, educators and interested lay persons. It contains much anecdotal and verbatim interview material that often provides fascinating excursions into the minds of the subjects.

Eiduson's work as reported here is so similar to that reported by Anne Roe in *The Making of a Scientist*, published ten years ago, and at that time the first of such studies to appear, that some comparisons are almost mandatory.

Eiduson's sample of 40 research scientists consisted of 16 biologists, 12 chemists, 6

geologists and 6 physicists, all of whom met the criterion of "choice of profession and its pursuit in academic and research-oriented atmosphere." They appear to range quite widely in research experience, effectiveness and achievement. Roe's sample consisted of 64 scientists nominated by colleagues as being outstanding in their fields of specialization. There were 22 physicists, 20 biologists, 14 psychologists and 8 archaeologists. Both investigators used the Rorschach and the Thematic Apperception Test plus open-ended depth interviews as the source of raw data. Roe, in addition, used a Verbal-Spatial-Mathematics test for an analysis of some of the cognitive abilities of her subjects. Due to the nature of her sample, Roe was able to find interesting and important differences between physical scientists and social scientists, and between theoretical physicists and experimental physicists. Eiduson's smaller sample of physical scientists did not permit such comparisons; her focus was somewhat more on the search for commonalities throughout, rather than differences within, her total group.

The major conclusions from these two studies are remarkably similar. Those given greater emphasis by both authors are:

- 1) Research scientists are similar to other human beings and can be understood within the same framework of personality theory.

- 2) They came from families that were predominantly middle class, with the father's occupation being either that of a business man or a professional.

- 3) There was little or no personal relationship between the son and his father. In some cases the father died during his son's early boyhood; in some he was absent from home for long periods of time, and in some he was a passive, dependent personality who made little or no impact on his family.

- 4) In their youth they commonly experienced periods of emotional and social isolation that forced them to fall back on their own inner resources for satisfaction. This led to the development of interests and activities of a solitary nature and a reliance on individual initiative in the solution of personal problems.

- 5) Early family relationships were usually cold and impersonal, with the subject forced either to look elsewhere for emotional support or to accept and endure considerable deprivation of his emotional needs.

- 6) As children, they gave evidence of intellectual capabilities far above the average. They quickly found adult approval and en-

couragement for intellectual achievement, which in turn served to reinforce further the satisfactions obtained from intellectual achievement.

- 7) There is no one specific personality type that is characteristic of the research scientist. For all practical purposes a wide variety of character types, defense mechanisms and modes of adjustment can be found among them.

- 8) Their primary emotional investment is in work and in people and things connected with it. Outside of work relationships they are comparatively uninterested in other people and do not relate to them easily. Typically, they maintain a large psychological and emotional distance between themselves and others.

- 9) They have outstanding intellectual capacities. Their thinking is objective, rational and logical and they are continually oriented toward the new and the novel. They strive to organize perceptions and experiences into new configurations and to formulate questions about the physical world in ways that are new and different. They are both willing and able to tolerate the anxiety that goes with ambiguity and uncertainty.

Eiduson elaborates more than Roe on the personality structure of her subjects and in the process provides data for a clearer understanding of their motivations. However, it is unfortunate that she tends to do this in the jargon of clinical psychopathology, which may reinforce the stereotype she wishes to destroy, namely, that scientists are always queer, eccentric or neurotic. In the chapter on the scientists' self-image she discusses some of the more practical aspects of the competitive rivalry among researchers. The inference is easily drawn that this is not far different from the infighting for status of industrial executives.

This book will undoubtedly help the educated layman and the interested educator to understand better the psychological nature of research scientists. At a time when so much attention is being focused on this socially important group, it is indeed fortunate that we have this kind of objective scientific reporting to add to Roe's earlier work. However, as far as the professional psychologist is concerned, neither of these studies seems to add much insight into the origins or dynamics of this very interesting group. It would seem that we have not yet discovered the key interactions between basic human attitudes and environmental ex-

periences that result in the new-born infant eventually developing into an effective research scientist.

JOHN R. WEIR
California Institute of Technology
Pasadena, California

Ginott, Haim G. *Group Psychotherapy with Children*. New York: McGraw-Hill, 1961, pp. 208.

Group play-therapy has met with increasing acceptance among practitioners in recent years. However, except for Slavson's extensive description of a specialized form of group therapy with children and some very limited references to the group approach by Axline and others, relatively little has heretofore been written on the subject. Announcement of the forthcoming issue of Ginott's *Group Psychotherapy with Children* was therefore greeted with enthusiasm by this reviewer who unfortunately did not foresee the hazards inherent in accepting a book title at face value. The reader, too, is apt to be disappointed if he expects from this volume some new or startling insights into group therapy with the younger client. To be sure, the author partially exonerates himself by prefacing his efforts with the explanation that much of the book's content applies equally to individual as to group therapy. Yet, even such preparation of the reader is insufficient to reshape the anticipation prompted by the inscription on the outer cover.

On the other hand, if this work be accepted simply as a treatise on play-therapy generally, then it is admittedly a very able, down-to-earth treatment of the subject which probably represents one of the best efforts since the advent of Frederick Allen's and possibly Virginia Axline's volumes a decade or two ago. Its focus is upon practical problems and their solutions to an extent not found in other books on child therapy. It deals with typical situations which a clinician inevitably meets with and attempts to focus upon the particular methods of handling them. Broad generalities are played down.

The opening chapter examines some of the objectives of group play-therapy within a psychoanalytically oriented framework, thereby setting the tenor for much that follows. Criteria are then formulated to provide a working basis for the selection and matching of prospective group members, with due regard nonetheless to the need for supporting research data. Thereafter the content is hardly distinguishable from what

might have been written on individual play-therapy. There is an excellent, if brief chapter on differential diagnosis utilizing play media in particular reference to the preschool child. In close conjunction therewith the author presents well-thought-out ideas relative to the stimulus-function of selected toys and play materials. The direction and progress of therapy is intimately tied to the behavioral-determinant nature of the toys themselves, each chosen with an eye to the child's central problem as well as to the case history.

Comments concerning some of the difficulties that arise in initiating therapy and in structuring the therapy relationship are devoted to preserving in the therapist a wholesome attitude with regard to the appropriate use of limits as well as a continued awareness of his own needs as they are likely to affect the child. To this end, the author prescribes a non-playing role for the therapist. Like Axline, he also advocates that structuring be postponed until the moment of need arises. There will be those who may take issue with what should or should not be said to the child at the outset. Thus, in contradistinction to Ginott it may be argued that limits implanted at the outset may well serve to reduce the likelihood of overwhelming anxiety, especially in the poorly controlled child.

The remaining pages discuss the person of the therapist, specific research problems, and the question of parent contacts. Reference to what constitutes an effective child therapist, rather than being a mere rehash concerning qualifications, prompts a degree of self-evaluation justifying the space accorded it by the author. The chapter on research in play-therapy, in keeping with the general trend of the text, has little to say about group therapy. With few exceptions, the studies cited have reference to individual play-therapy, their application to investigations of group phenomena being a matter for the reader to speculate upon. True, the suggestions for proposed study into the nature of the play-therapy process and its relative effectiveness as compared with other treatment methods are not to be discounted. But it is almost as though a clear declaration of research interest has become the *sine qua non* of respectability, without which any pronouncements based upon mere clinical experience border on heresy.

In discussing the parents' role, the author expounds rather convincingly on the economics of group screening as a tried pre-intake method designed to offset the disadvantages

of a waiting period and to render the whole intake procedure more effective. Insofar as collateral group treatment of the parent is concerned, some will not agree altogether with the distinction made between group therapy, group counseling, and group guidance. But they can hardly deny the importance of working with the parent at his current level of readiness, be it in relation to intra-personal or everyday reality problems or simply a need for help in understanding the facts of child growth and in developing appropriate ways of responding to them. The point is aptly made that too often the parent is not accepted where and for what he is at the time, but instead is browbeaten into adopting a patient role when he is not ready and may not even need such treatment. All that he may require is information. This becomes the function of group guidance which, pending further study, is offered as the most promising means of meeting the needs of the maximum number.

This is a concise, compact volume written in a lucid, simple, relatively non-technical style which makes for easy reading. It is not laden with endless case illustrations or descriptive accounts of isolated therapy sessions. For the most part, the author presents his own viewpoint or position and does not leave the reader to guess. He takes a stand rather than merely presenting or attempting to summarize different therapeutic approaches. Unfortunately, a void still remains insofar as the treatment of the inpatient child is concerned and, for that matter, many aspects of group psychotherapy with children. Mention could have been made of the two-member group as well as of the possibilities attending the pre-adolescent non-activity group. And what of the role of the co-therapist? Whatever the book's shortcomings, it conveys so much of the author's sensitivity and keen perceptiveness as to render it of inestimable value to the new clinician and worthy of perusal by the more sophisticated practitioner.

GORDON FILMER-BENNETT
Winnebago State Hospital
Winnebago, Wisconsin

Holtzman, Wayne H., Thorpe, Joseph S., Swartz, Jon D., & Herron, E. Wayne
Inkblot perception and personality.
Austin: University of Texas Press,
1961. 417 pp. \$8.00.

culmination of an extensive program of research inquiry into the development of a psychometrically sophisticated alternative to the Rorschach Method. The resultant of this inquiry is a scholarly treatise, providing a cogent analysis of psychometric weaknesses of the Rorschach Method and offering an alternative approach to inkblot perception that circumvents these psychometric inadequacies while retaining the unique projective approach to personality assessment that the Rorschach Method offers. Those who are aware of the often caustic criticisms that have been leveled at inkblot techniques from the standpoint of their psychometric imperfections will welcome the advent of an approach to inkblot techniques that promises to obviate the greater majority of these difficulties.

The Holtzman Technique, consisting of alternate parallel forms of forty-five inkblots each, provides for a single response to each blot, thus obviating the problem of variations in response total that has been a major source of annoyance to psychometric critics of the Rorschach Method. In addition, scoring criteria for the response variables have been extensively refined to lessen ambiguities arising from the requirements of subjective judgment, and the total series of response variables has been factor analyzed to define underlying congruences thought to represent certain personality attributes. The refinement of scoring criteria contributes to interobserver reliability while the analytic scrutiny of the response variables provides for an approach to an interpretative analysis that is grounded upon consistent and clearly definable terms that tend to remain relatively constant across individuals within particular groupings. An adequate body of normative data is presented that should allow other investigators an opportunity to explore the validity generalization of the Holtzman Technique across similar populations.

Since the Holtzman Technique has not been widely available prior to this time, little opportunity has been afforded for extended clinical usage. However, research inquiries thus far reported seem encouraging and it is to be expected that further clinical research will be forthcoming. In particular, the psychometric soundness of the technique will encourage clinicians to explore its research possibilities and the provision for translating from Holtzman variables to Rorschach variables ought to encourage comparisons of the clinical utility of both instruments. The limited comparative research already accom-

degree of commonality for both techniques and it is probable that the usefulness of each technique can be enhanced by further research into their mutual contributions one to another.

With respect to other considerations generated by a review of the present work it should perhaps be noted that many clinicians have considered that the Rorschach is in reality two separate techniques; first, the truly projective aspects of an individual's spontaneous associations to such inkblot materials; and, second, the more formalized interrogative aspects of the inquiry for the determination of scoring categories. Thus, the Rorschach has had both projective as well as psychometric aspects and it has been in the latter area that the majority of criticisms have arisen. The Holtzman Technique by virtue of its basic psychometric soundness should offer the possibility of validating, on an experimentally demonstrable basis, certain Rorschach findings that have heretofore been equivocal because of the statistical inadequacies of the Rorschach Method. Similarly, as a body of knowledge accrues around the Holtzman Technique, it should become apparent from what areas the Holtzman Technique can be implemented by drawing upon the Rorschach literature. Such mutual enrichment will represent a distinct advance in the establishment of sound foundations for the assessment of personality through projective methods.

A. BARCLAY, PH.D.
Saint Louis University and
Cardinal Glennon Memorial
Hospital for Children

Ittelson, William H., Kutash, Samuel B., ed: *Perceptual Changes in Psychopathology*. New Brunswick: Rutgers University Press, 1961. 262 pp. \$9.00.

Perceptual Change in Psychopathology is a progress report on research conducted at the East Orange Veterans Administration Hospital. The underlying assumption of this research is:

Much of psychopathology can be conceptualized as distortions either of perceptual content or of perceptual processes, and that the course of psychotherapy can be conceptualized as a course of perceptual change.

The initial two chapters by the editors present an elaboration of this theoretical framework. This involves definition of terms such

as "perception" and "personality" in the perceptual orientation. These chapters present a good foundation for the research material which follows. The remaining chapters are research studies conducted by numerous individuals, many for dissertation research.

Perhaps a word about dissertation research is in order. Much of the dissertation research done in many clinical programs is very anxiety-laden because of the pressure for excellent scientific form. This produces a dissertation which, like Minerva, springs apparently exquisitely complete from the brow of the author. The struggle for understanding the meaning of data in relation to hypotheses is often not apparent. To manipulate data by combining groups, or re-examining this or that part of the data becomes not quite respectable because it was not initially proposed in the procedure. Such a restrictive approach to research may discourage many psychologists from ever again attempting research.

Perceptual Change is valuable not only because of the research it presents, but because of the manner in which it is presented. The problems of the research and the researcher are evident and the solutions developed to deal with them are offered. The studies are admittedly not exactly as the author would have liked. Thus, there is an alive and energetic feeling about the research which may encourage others to embark on the research process.

The research presented is good from many points of view. It is clearly related to the framework of the relationship of perception in personality and perceptual change in psychotherapy. The studies are neither too broad to have meaning, nor too limited to have value. The procedures often involve the Ames techniques of the "thereness-thatness" demonstration, a number of applications of aniseikonic lenses, the leaf room, etc. The aniseikonic lenses are a favorite for producing perceptual distortion and other techniques were developed for specific studies. The independent variables include such varied concepts as personality orientation, nature of group interaction, stress, and anxiety. There is one study relating personality, skeletal-muscular, and perceptual functioning. There are studies on perceptual change related to self-percept, self-products, and ultimately to the therapeutic process. Thus the concept of perceptual change is applied in a wide variety of areas and always with meaning.

Because "perception is a process carried

out in that pattern of processes called personality," the study of perception is important in its broadest sense. This book clearly fits in the movement toward understanding personality and the psychotherapeutic process through understanding of perception. The ideas suggested by this book in the areas of perceptual development and functioning in children, perceptual change in psychotherapy, and perceptual hypotheses about projective testing are exciting.

DONALD L. WESTON
Univ. of Maryland
Medical School

Levitt, E. E. *Clinical Research Design and Analysis in the Behavioral Sciences*. Springfield, Illinois: Charles C. Thomas, 1961, pp. 199—xxii. \$8.50.

Thank you, Dr. Levitt, for your efforts to make the aims and methods of science intelligible to the clinician. Books such as this one are badly needed. It is a refreshing contrast to the usual, involved statistical analysis texts in which there is "... no attempt made to present methods in an easily comprehensible form."

The book is rather brief and at no point wordy. Style is clear and interesting. Coverage is limited by the author's goals but is fairly adequate. In a well organized fashion he progresses from philosophy of science through design, execution, statistical analysis, drawing inferences, to the final task of writing the report.

Because of the laudable aims of the book I had hoped to begin the review with "Bravo, Dr. Levitt!" However, reading the first chapter, a philosophy for the behavioral scientist, quickly dissuaded me. This chapter discloses a lot more about the author than do his *preface* and *introduction*.

Dr. Levitt seems to be a philosophical eclectic; tough-minded empiricism superimposed on a tender-minded clinical approach. He cites the logical positivist, Professor Herbert Feigl, as an important influence, yet employs positivism whimsically. While the absence of a logically consistent philosophy of science obliges me to warn students against uncritical reading of the first chapter, the case is somewhat different for the clinician, to whom the book is directed. Any absorption of the philosophy underlying empirical science will help the clinician design and interpret research more intelligently. Since

mend their reading of the book in its entirety.

With a sincere desire to encourage the author's efforts with a favorable review I hoped that the first chapter would be the only weak one. Alas, errors and logical inconsistencies exist in almost every chapter. It is hard to decide whether these, let us call them weak points, arise from the author's deliberate oversimplification or from a few areas of naivete. While I fear the latter, we shall give him the benefit of the doubt and say that he is courageous to publish such oversimplification.

Dr. Levitt is quite right in feeling that simplification is necessary to make important material comprehensible to the research neophyte. So, despite its weak points, the book should prove of great value to many clinicians. Regretfully, I cannot recommend it highly to psychology students.

It would be interesting and valuable to be able to discuss what I feel are the weak points in the presentation. Such a review would offer the author criticisms to refute, and I have no doubt he could offer persuasive defense on many of them. Then, too, the reader should be warned not to accept certain points uncritically in case he studies from the book. But the reader of the book often will lack the background to understand the more complex issues expressed in a contrast of opinions. At any rate, the points of issue are too many and too scattered to permit coverage in a brief review.

By way of illustration, we will look at the type of problem that arises when simplification is attempted. Dr. Levitt substitutes "co-fluctuate" for the more usual term *covary*, when discussing correlation. His motive seems to be that of simplifying language and improving readability. But in using fluctuate rather than vary there is an implication that the variation in subjects' scores is chance rather than systematically related to characteristics of the individuals. So with good intentions Levitt avoids Scylla only to be wrecked by Charybdis.

Conceptions and applications of the terms, construct and population, bothered me most because unnecessary complexity seemed to be introduced. There are a few outright errors of minor importance, but it takes little to set statisticians' teeth on edge. One might guess that this review will be more favorable than most, especially if the others are written by statisticians.

Old "pros" will find nothing new statistically in the book. The author shows consid-

it to style and organization in order to accomplish his teaching goals. While it is not a scholarly book, the author has worked hard to achieve his goals with some resulting degree of success. While I cannot recommend it highly to students it deserves the clinician's attention. If most clinicians mastered the knowledge in this small book, psychiatry and clinical psychology would appear less anachronistic in the space age.

WILSON H. GUERTIN
University of Florida

Lindzey, Gardner. *Projective Techniques & Cross-Cultural Research.* Published by Appleton, Century-Crofts, Inc., N.Y., 1961, 339 pp., \$6.00.

Clifton, James A. and Levine, David. *Klamath Personalities: Ten Rorschach Case Studies.* University of Oregon Press, Eugene, 1961, 80 pp.

One can perhaps safely say that Lindzey's volume is one of the few books, and possibly the first one, that attempts to compile and assess studies on pre-literate societies employing projective techniques. In a sense, this book is an endeavor to combine the area of anthropology and psychology. Although the bulk of the volume deals with the anthropological studies, it is a little more ambitious than a mere compilation. It delves into the history (although briefly) of each projective test, attempts to define and classify projective tests, tries to lay bare the usual assumptions that are made or involved in the interpretation of any projective test (especially the Rorschach), and finally attempts to assess the past contributions of projective tests in anthropological studies and what the future role of the projective tests may be in such studies.

The author, in true scientific spirit, goes on to define what constitutes a projective test. He maintains that there are two criteria for defining a projective test — primary and secondary. In terms of primary criteria, he defines the projective test as "an instrument that is considered especially sensitive to covert or unconscious aspect of the behavior, permits or encourages a wide variety of subject responses, is highly multidimensional, evokes unusually rich or diffuse response data with a minimum of subject awareness concerning the purpose of the test". For the secondary criteria, he continues, "the stimulus material presented by the projective test is

ambiguous, the interpretation of the test depends upon holistic analysis, the test evokes fantasy responses, and there are no correct or incorrect responses to the test". As far as classification of the projective tests is concerned, the classification based upon differences in the type of responses is easily the most important and the one that should be most emphasized. By type of responses he means whether the test evokes association, construction, completion, etc.

In the chapter "What are the Theoretical Foundations of Projective Techniques", the author bemoans the fact that there are hardly any theoretical foundations of the tests. He seems to blame the tests for the fact that one cannot make any definite or concrete prediction from them or at least the studies have not thus far shown it to be so. Besides, there are many assumptions involved in making an interpretation, and interpreters themselves rarely agree with each other. Throughout the chapter one gets the impression that Lindzey talks as if there is a distinct difference between personality theory found in general psychology and personality theory derived from projective tests. Further, he implies that theories of personality can lead to better predictions than the projective tests because the tests lag behind in precision, scientificness, etc. It is obvious to every worker in this field that the tests can enable no better prediction than the state of general knowledge or theory in any area. Therefore, it seems to the reviewer that not only has the author chosen a wrong target for his attack in blaming the tests alone for the poor state of affairs, but has overlooked the far more critical problem of the still inadequate level of sophistication of personality theory. In this sense, it may be said that he has created a straw man.

His recommendation is that those working with projective tests should try to develop 'miniature theories' rather than trying to develop an all encompassing theory of personality based upon projective tests. For example, what is M (Human movement on Rorschach) supposed to mean and why. According to him, it is such an approach which is likely to prove more fruitful.

About one-third of the book is devoted to reporting cross-cultural studies and critiques thereof. He classifies the studies done on pre-literate societies in three sections, namely, national character or modal personality studies, acculturation studies, and procedural studies. The author systematically discusses the shortcomings and the contributions

of each of these studies. He summarizes his criticisms as follows: that the interpreter had a knowledge of the culture; failure to provide a free description of the circumstances under which the test is administered; general failure to explore the possible contribution of a non-personality factor (or the role of situational factors); the probable meaning of the test-taking activity; the examiner's influence upon the performance; a failure to select for comparison groups; mechanical application of the scoring system and interpretive generalizations; a tendency to take group averages and treat them as descriptive of the group as a whole; and finally, relatively little evidence of cumulation of sophistication and wisdom on the basis of all these studies.

Apparently Lindzey criticizes the studies from the standpoint of a 'pure' science, in that the non-personality or situational factors should be controlled, the groups should be matched, etc. It is known to everyone that even in America such stringent scientific requirements are almost impossible to achieve. Therefore, one may say that the author expects a perfect study in a field, namely cross-cultural research, which is at best pre-scientific. In the same vein, the author compares these studies to a theoretical ideal or utopia which on the face of it is not possible to achieve; e.g.; Lindzey wants the field worker to spell out the test-taking attitude. This attitude is usually so covert that if a field worker is able to delineate, describe and control such covert attitudes, he really does not have to study the society, since he would already know so much about it. Hence, one may say that Lindzey criticises the performance of these practical studies from a purely theoretical-idealistic viewpoint and similarly evaluates an area of pre-scientific knowledge from a 'pure' scientific standard.

To this reviewer, what struck a warning note, which Lindzey seems to overlook, is a certain ethnocentric approach of researcher or field worker. Most of the societies reported were more neurotic, more conflict-ridden, more anxious, more repressive, less spontaneous, etc., as compared to America. Other studies done on the same or similar societies do not quite support such findings. In view of the large number of such findings, one may well suspect a consistent ethnocentric bias of the investigators.

Perhaps at this stage of our knowledge, and in such a field of study one cannot, by any

Lindzey. In the opinion of this writer, the only way in which we can use projective tests effectively and fruitfully in other cultures, is to rely upon the interpretation and the judgments of the individuals who are thoroughly familiar with particular cultures and at the same time are experts on projective tests. We have simply to take their formulations, their interpretations, as the starting hypothesis and from there on verify or disprove whatever these people say. Otherwise, one is liable to be lost in ethnocentric bias, in insignificant details and in fastidiousness of methodology, without getting to the psychological reality of the culture.

In spite of what has been mentioned, the book is indeed a refreshing one to read. It is well-documented, sharply written and it covers practically all that has been done in this area in the past decade and a half. It tries to show both the strong and weak points of each study and should provide stimulating and thought-provoking, although somewhat pessimistic, reading for any graduate class. Regardless of what has been said of Lindzey's criticisms of the studies, such criticisms do play a positive role, for they show the pitfalls that one should avoid and point the direction in which one should go.

Clifton and Levine's mimeographed work which is expanded from a portion of a Ph.D. dissertation, deals with Klamath Personality. The Klamath is a Red Indian tribe residing in the northwest corner of the United States—the states of Oregon and Washington. In 1955 the tribe had 2118 members and of these 40% live away from the reservation. Today, more than 70% of the tribe members are genetically less than half Indian. Based upon a series of criteria—education, length of stay in reservation, etc.—the Rorschach, the TAT and the life history of 54 people were obtained. The ten cases presented here were selected from the original sample so as to provide as much variety as possible.

The aim of this study is to record variant attributes of personality, understanding of the psychological, especially the interpersonal aspects of Klamath personality, and the synthesis of the psychological and cultural data which may throw additional light on psychological consequences of acculturation. The Rorschach test was used as a means of gaining insight about the adaptive mechanisms of the individuals in a particular sub-group undergoing the acculturation process and it was interpreted blindly. The author (Levine) states that in his interpretations of Rorschach data, which he compares with Piaget's

method of interpretation.

From the Rorschach test, he concludes that the Klamath subjects have, in general, a "cautious guardedness" in their interpersonal relations and that women are better put together in that culture. Both of these conclusions, independently arrived at by the Rorschacher, appear to be confirmed by the cultural study. It is the impression of the reviewer that had the psychologist used more of an eclectic approach to the interpretation of the Rorschach records, rather than sticking so closely to Piotrowski's method, he might have been able to pick up many other facets of the Klamath personality structure and their interpersonal relationships, e.g., the role of aggression and passivity.

The present study, contrary to certain prevailing beliefs, shows that the more the Rorschach interpreter is familiar with the culture that is being studied, the better the chances of his arriving at meaningful interpretations of the records. It is, therefore, unfortunate that the interpreter, although extremely familiar with Klamath culture, did not make much use of the content of the

records. To this reviewer, it appears that the use and the interpretation of the test contents provides one of the more reliable guides to the subjects' acculturation level. Although no comparison within the same culture is made, the records of these ten Klamath cases are compared with the other Red Indian tribes of the United States who are going through a similar process of acculturation. Thus, the study shows clearly the strains and effect of acculturation.

Altogether, in the opinion of this reviewer, projective tests have proved useful in cross-cultural studies. There are many instances where they have not proved so fruitful because of a relative lack of familiarity of the investigator either with tests or with culture or both. However, let us hope that the future investigators will be able to benefit from past errors — many of them mentioned in these volumes — and provide us with more scientific and psychologically meaningful studies.

SOHAN LAL SHARMA, Ph.D.
Los Angeles Psychiatric Service

ANNOUNCEMENTS

SEVENTH CONGRESS OF THE INTERAMERICAN SOCIETY OF PSYCHOLOGY

The five-day meeting of the Seventh Congress of the Interamerican Society of Psychology was held in Mexico City from December 19 to 23, 1961. Approximately 650 delegates registered for the Congress. There were representatives from 11 countries: Argentina, Brazil, Chile, Columbia, Cuba, Guatemala, Mexico, Panama, Peru, the United States (including Alaska), and Venezuela.

The following were the official sponsors of the Congress: Social Security Institute of Mexico, National University of Mexico, Rama Mexicana de la SIP, the Centro Investigaciones Sociales (CISAC). Sessions were held at the Centro Medico of the Seguro Social, recently built, which has the finest facilities for professional congresses in all North America. A great degree of interaction and exchange was made possible by the availability of trilingual translating facilities for English, Spanish, and Portuguese. The Organizing Committee was under the able direction of Dr. Guillermo Davila G. of Mexico City.

The major themes of the meeting included Culture and Personality, Experimental Psychology, Applied Psychology (Educational and Industrial), and Psychology and Mental Health. Ninety-five papers were presented. Five plenary meetings were devoted to major papers given by Dr. Gustave Gilbert (U.S.) on "The Needs for a Comprehensive Bio-Social Theory of Personality"; Dr. Erich Fromm (Mexico), on "The Revolutionary Character Structure"; Dr. Carlos Alberto Seguin (Peru), "Language and Communication"; Dr. Abraham Maslow (U.S.), "The Scientific Study of Values"; and Dr. Robert B. Malmo (Canada), "Contribution of Experimental Psychology of the Clinical Interview."

A Round Table was held concurrently with the Congress, organized by the CISAC with Dr. Wayne H. Holtzman (U.S.) and Dr. Rogelio Diaz-Guerrero (Mexico), as co-chairmen on the subject "Culture Shock and Social Change." With the support of the United States Agency for International Development, a number of distinguished Latin American social scientists were invited, primarily anthropologists,

psychologists, social psychiatrists, and sociologists. Another Round Table series on Group Psychotherapy was chaired by Dr. George Bach (U.S.).

The program also included two symposia on "The Development of the Personnel Subsystem in Industrial and Military Machine Systems", with Dr. Jerome Ely (U.S.) as coordinator, and "Transcultural Impact at the U.S.-Mexican Border", with Dr. Rogelio Diaz-Guerrero as coordinator.

Our Mexican hosts provided an excellent program of social activities, including a gala night of entertainment by the National Dancers of the Seguro Social.

Arrangements have been made to publish, in English and Spanish, the proceedings of the entire Seventh Congress through the financial planning of the Mexican Branch, under the leadership of Dr. Guillermo Davila G.

The officers of the Society designated for 1962-63 were Dr. José Bustamante, President; Dr. Gustave M. Gilbert, Past-President; Dr. Harold H. Anderson, President-Elect; Dr. Victor D. Sanua, Executive Secretary for North America; Dr. Rogelio Diaz-Guerrero, Executive Secretary for Mexico and the Caribbean Islands; Dr. Fernanda Monasterio, Executive Secretary for South America; and Dr. Rafael Nunez, Treasurer. Additional members of the Society's governing board are the following Vice-Presidents: Dr. Wayne H. Holtzman for North America, and Dr. Carlos Alberto Seguin for South America.

The Eighth Congress of The Interamerican Society of Psychology is scheduled for Easter of 1963 and will be held at Mar La Plata, Argentina, under the sponsorship of La Plata University.

VICTOR D. SANUA, PH.D.

1962 ANNUAL WORKSHOP IN PROJECTIVE DRAWINGS

The 1962 Annual Workshop in Projective Drawings will be conducted at New York State Psychiatric Institute, New York City, this year by Emanuel F. Hammer, Ph.D. and Miss Selma Landisberg, M.A., on July 23-26, from 10:00-12:00 and 1:00-3:00 daily. The workshop will provide a grounding in fundamentals and then go on to advanced considerations of differential diagnosis, psychody-

namic appraisal, anxiety, the individual's psychological resources as treatment potentials, and the use of projective drawings in therapy. In addition to the H-T-P and Draw-A-Person Test, the workshop will include the Draw-A-Family procedure, the Unpleasant Concept Test, the Drawing Completion Test, the Draw-An-Animal technique, and free doodles.

The Clinical Application of Projective Drawings, Charles Thomas, Publisher, East Lawrence Avenue, Springfield, Illinois, is recommended as advanced preparation for the Workshop. For information as to admission, fees or requirements, write to Miss Selma Landisberg, 166 East 35th Street, New York City.

REGIONAL REPORT

Mrs. Patricia M. Bricklin was elected president of the Philadelphia Division of the Society for Projective Techniques. Dr. Barry Bricklin was elected secretary-treasurer (1961-62). The fact that the treasury is a bit low has to do with the fact that Philadelphians are recalcitrant about paying dues and not with the fact that both executive officers are from the same family. You must believe this! Please!

The program this year has included a talk by Dr. Barry Bricklin and Dr. Zygmunt A. Piotrowski on the Hand Test, a new projective test which predicts aggressive overt behavior, and a talk by Dr. Fred Brown on the planning and evaluation of psychotherapy by means of projective tests. Another meeting will be announced for late spring.

Clearance Sale on Back Volumes

The Journal of Projective Techniques has a limited number of copies of each of the volumes published from 1947 to the present. For a short time we are willing to dispose of them at cost to individuals and *below* cost to libraries and other institutions (U. S. and foreign).

ORDER BLANK

Society for Projective Techniques, Inc.
Adelphi College
Garden City, N. Y.

Enclosed find \$_____ for the following order:

Bill me _____ for the following order:

Volumes	Year	Institution \$4.00	Individual \$5.00
11 & 12	1947-48	_____	_____
13	1949	_____	_____
14	1950	_____	_____
15	1951	_____	_____
16	1952	_____	_____
17	1953	_____	_____
18	1954	_____	_____
19	1955	_____	_____
20	1956	_____	_____
		\$4.50	
21	1957	_____	_____
22	1958	_____	_____
23	1959	_____	_____
24	1960	_____	_____

Name or Institution: _____

Street _____

City _____ Zone _____ State or Country _____

JOURNAL OF PROJECTIVE TECHNIQUES

VOL. 26

SEPTEMBER, 1962

No. 3

. . .

EDITOR

BRUNO KLOPPER

EXECUTIVE EDITOR

BERTRAM R. FORER

ASSOCIATE EXECUTIVE EDITORS

WALTHER JOEL
JEANNE M. REITZELL
EDWIN S. SHNEIDMAN

STAFF EDITORS

EVELYN CRUMPTON
FLORENCE DIAMOND
NORMAN L. FARBEROW
HERMAN FEIFEL
FRED J. GOLDSTEIN
JAMES HOLT

WALTER G. KLOPPER
LEONARD KOROT
JEROME REISEL
NICHOLAS ROSE
STANLEY F. SCHNEIDER
JOHN WEIR

IRLA LEE ZIMMERMAN

A Refutation of Eriksen's Sensitization-Defense Hypotheses¹

STANLEY ABRAMS
Camarillo State Hospital, California

There is an increasing body of research that appears consistent with the concept of repression. Of interest, however, has been the observation that repressive behavior is not "unipresent" or even normally distributed, but rather, conforms to a bimodal distribution. Eriksen (1951) has labeled the two such groups as sensitizers and defenders. The former dwell on their failures, while the defenders emphasize their successes and forget their failures. Rosenzweig (1938) has also discovered this dichotomy and demonstrated marked personality differences between the two groups. Those Ss who were better able to recall successes than failures tended to gloss over their failures and forget them. In contrast to this the Ss who recalled failures tended to blame themselves or others for these failures. Similar results were attained by Mainord (1956) who discovered that the sensitizers tended to cope with their failures while the defenders used repression. In another investigation Eriksen (1952a) has shown that the sensitizer learns affective words as easily as neutral ones, but defenders require more trials to learn and relearn these affective words.

Eriksen (1952b) has hypothesized that these personality differences are constant, regardless of the psychic functions involved or the stimuli acting on the individual. He demonstrated that with a prior knowledge of the effect of threat on memory he was able to predict the individual's response to perceptually threatening stimuli. Those who forgot memory

failures tended to show a similar defense in perception. In an investigation by Rosenstock (1951), a positive relationship was found between memory and perception. Repressive behavior was reflected in both recall and perceptual tasks. Eriksen and Browne (1956) discovered that needs which were acceptable received little or no inhibition, while unacceptable needs were kept from consciousness. Among openly aggressive and homosexual Ss no perceptual defense occurred. These individuals were able to recognize significant pictures as quickly as neutral ones. The authors concluded that needs may influence perception either as sensitizing the individual or acting as a perceptual defense.

Other observations have indicated that the sensitization-defense dichotomy also occurs in projective testing. Lazarus, Eriksen and Fonda (1951) attained a reliable correlation between auditory perceptual recognition of sexual, aggressive and neutral sentences and performance on a sentence completion test. These authors suggest that it might even be possible to identify unconscious determinants of behavior through the use of a perceptual defense technique. Along this line Eriksen (1951) presented evidence consistent with the notion that the TAT and Rorschach reveal only non-repressed needs, while weak or repressed needs are not manifested. Substantiating results have been attained by Eriksen and Lazarus (1952). In a similar investigation, Murstein (1956) found that Ss who were consciously hostile tended to project more hostility on the Rorschach than either hostile non-insightful or non-hostile Ss, suggesting a relationship between awareness and sensi-

¹ This study is based on a doctoral dissertation done at Temple University. The author wishes to express his appreciation to Drs. James D. Page and Horace A. Page for their many contributions.

tization, and repression and defensiveness.

PROBLEM

The purpose of this study is to test Eriksen's sensitization-defense hypotheses, to determine whether this dichotomy exists in the area of hostility and whether it is constant for various psychic functions. Hostility was chosen as the trait to be studied because it is present to some extent in all individuals, but more important, because it is subject to prohibition in our society and would consequently, be considered highly susceptible to "repression". It was hypothesized that those individuals who were hostile and aware of it would respond as sensitizers, while those who were hostile and unaware of it would respond as defenders.

Eriksen's constancy hypothesis was tested by comparing three areas of psychic functioning—recall, projection and preference. It was expected that an individual giving a sensitization or defensive response in one of the three areas would tend to respond in the same direction in the other two. Thus, a sensitizer would recall, project, and prefer more hostile stimuli than a defender.

METHOD

Subjects

The sample from which the Ss were drawn consisted of 250 student nurses and oral hygienists. All were female and approximately 19 years of age. These particular groups were selected because peer ratings were to be used as the criteria for hostility and it was assumed that their frequent close contact in work, school and in the dormitories would allow them to accurately judge their classmates in this respect. The first step involved identifying girls who were well enough acquainted to rate one another. For this purpose each of the 250 Ss was given a list of the names of all the girls in her class and instructed to place a check mark beside the names of those girls whom she knew well enough to make

an accurate personality description of. From the returned lists, paired comparisons were constructed. Each S was given 5 other Ss to compare to each other and to herself in regard to the degree of hostility possessed. The 5 Ss compared by each rater were taken from the lists of girls each S had indicated she knew well. Ratings were obtained by giving each rater an envelope containing 15 cards bearing the names of 2 Ss. Each S appeared on 5 different cards, each time being paired with a different S. Since each S was compared to 5 other Ss by 5 different raters, an hostility rating ranging from 0 to 25 was possible. However, to avoid any possible biasing effect produced by the rater comparing the S with herself, the self-other ratings were not used here. Thus, each S could attain a possible hostility score from 0 to 20.

Self rating scores were determined by totaling the number of times the rater indicated that she was more hostile than a S with whom she compared herself. Since she was paired with 5 Ss, she had a possible hostile self rating ranging from 0 to 5. Frequency distributions on both self and other ratings were constructed for all 250 S. Arbitrary percentile levels were selected to differentiate between the high, average and low hostility ratings. Individuals falling in the upper 20th percentile were scored as high hostility while those in the lower 20th percentile were considered as low in hostility. Those Ss between the 40th and 60th percentiles were rated as average in hostility. From the results of the paired comparisons and the frequency distributions the following three experimental groups were chosen:

1. The twenty Ss who were rated as most hostile and who rated themselves in this way. They attained a "self rating" of 4 or 5 and an "other rating" of 16 to 20.
2. The twenty Ss who were rated as most hostile but who did not rate themselves in this way. They attained

a "self rating" of 0 or 1 and an "other rating" of 16 to 20.

3. The twenty Ss who were rated as being average in hostility and who rated themselves in this way. They attained a "self rating" of 3 and an "other rating" of 9 to 11.

Materials

An opaque projector and screen were used to present the following techniques:

1. Rorschach Ink Blot Test
2. Recall Test
3. Preference Test

The recall and preference tests were constructed because no suitable techniques were available to measure these functions and their relationship to hostility. They were developed in the following manner: from a series of symbols found to be representative of hostility by Schafer (1954) and Phillips and Smith (1953), 75 were selected for use in the preference and recall tests. Another 75 symbols which appeared to be free of hostile connotations were added to this. Each symbol was then drawn on an 8 by 11 inch sheet of paper and were shown by means of an opaque projector to a second sample of 60 college students. This group was composed of both males and females who were taking courses in Introductory Psychology. The Ss were instructed to free associate to each picture, writing down whatever occurred to them during the 20 seconds which they viewed each picture.

The responses were then evaluated as hostile or non-hostile. A response was considered hostile if the S used words such as: "kill", "hate", "maim", etc. Nothing which might only indirectly suggest hostility was accepted as an hostile response. Responses which contained no hostile words were viewed as non-hostile. A picture was accepted as being symbolic of hostility if at least 75 percent of the group responded to it with some hostile term. Those pictures to which no aggressive reaction was given by any S was scored as a non-hostile symbol.

The pictures which did not fall into either the hostile or non-hostile category were discarded. The remaining 105 symbols were shown to 60 volunteers from the original sample of 250 for the purpose of attaining a preference rating for each symbol. The pictures were randomly distributed and shown 5 at a time to groups of 10 Ss. Each S was to choose the picture she preferred most of the 5 that were shown. This was continued until all 105 pictures were shown. The symbols were then reshuffled and shown to the next group of 10 Ss. Preference values were assigned to each picture by allotting 1 point to a symbol each time it was chosen as "most preferred" of the group of 5 with which it was presented. Each picture was then photographed individually so that it would be of approximately equal size and color contrast as the other symbols. Hostile symbols were then paired with non-hostile symbols of equal preference value. From these symbols, 4 pair were arbitrarily chosen for each of the 8 cards. Thus, each of the 8 cards was composed of 4 hostile and 4 non-hostile symbols. The symbols were placed in 2 rows, one above the other, with 4 symbols in each row. Each hostile symbol alternated in position with each of the non-hostile. The first 4 cards were used as the test of recall and the second 4 as a measure of preference.

Procedure

All the tests were administered in one sitting to groups ranging from 20 to 50 Ss. Each test was presented by means of an opaque projector and the Ss indicated their responses on a prepared answer sheet. All of the original sample of 250 Ss were administered all the tests and the records of the 60 Ss comprising the 3 experimental groups were later identified and analyzed.

In the test of recall, each card was shown for three seconds, after which the S was given time to write the names of the symbols she recalled.

The same procedure was followed for all four cards of the recall test. The individual's recall score was the percentage of hostile symbols remembered as compared with the total number of symbols recalled. The remaining four of the eight cards prepared were next administered to the groups. This time the cards were projected on the screen for three minutes and each S was requested to list the two symbols she liked the most and the two she liked the least on each card. The individual's preference score was determined by algebraically subtracting the total number of hostile symbols preferred from those disliked. Immediately following the completion of the preference test, the Rorschach was presented. The S was instructed to give three responses to each of the ten cards. A score was ascertained by giving one point for every hostile response given. A response was considered hostile if it were similar to those hostile responses listed by Phillips and Smith (1953) or by Schafer (1954).

RESULTS

The findings indicate that the predicted differences were substantiated in only some cases. As shown in Table

TABLE I—Chi Square Tests of Differences

Subjects	Preference	Recall
Insightful ^a —Repressed	4.90**	.401
Insightful ^a —Average	.109	.002
Average ^a —Repressed	3.61*	.401

I, the results for the functions of recall and preference were in the predicted direction for all three hostility groups. That is, the hostile insightful Ss preferred and recalled more hostile symbols than either the average or hostile repressed groups. The average hostility group in turn, preferred and recalled more hostile stimuli than the hostile repressed Ss. However, statistical significance was attained only in the preference test where the insightful hostile Ss preferred more hostile

symbols than the repressed group. Although the average group preferred more hostile stimuli than the repressed Ss, only the .10 level was reached. In the case of recall, no significant differences were found among the three groups.

TABLE II—Chi Square Tests of Differences

Subjects	Condition Projection
Insightful ^a —Repressed	.109
Insightful ^a —Average	12.5***
Repressed ^a —Average	4.95**
^a Higher	
* $P < .10$	
** $P < .05$	
*** $P < .001$	

Table II shows that in projection, in contrast to Eriksen's hypothesis, both hostile groups gave significantly more hostile responses to the Rorschach than the average group. The hostile insightful Ss did, however, project more hostile symbolism onto the ink blots than the repressed group, but statistical significance was not attained.

A Spearman Rho was used to determine if the Ss responded in a constant manner to the hostile stimuli, regardless of which of the three psychic functions were involved. Although the insightful Ss as a group were always more responsive to the hostile stimuli than the other groups, no significant demonstration of constancy was attained. All of the correlations were low and none was statistically significant.

These findings would cast some doubt on Eriksen's sensitization-defense hypotheses within the area of hostility. It should be mentioned, however, that the results may have been affected by the narrowness and homogeneity of the sample. Student nurses and oral hygienists may even have been drawn to these particular vocations as a means of dealing with their own hostility. Thus, generalization to other groups is somewhat limited.

DISCUSSION

Although Eriksen's constancy hypothesis was not substantiated, it was found that in some cases a bi-modal distribution of sensitizers and defenders did occur.

In the area of preference, the results indicate that individuals who are hostile and aware of it are inclined to prefer more hostile symbols than those Ss who repress this trait. The insightful group, thus, tends to respond to threat with an increased awareness and response to the threatening stimuli, thereby coping with threat through this increased awareness. In a sense, knowing what and where the threat is, is in itself anxiety reducing. The defenders, in contrast to this, reduce threat with what appears to be suppression. They consciously reject that which is unpleasant and select instead, the more pleasing. This reaction appears to be similar to the one found in many perceptual defense investigations. Here, rather than an unconscious defensive process occurring, the individual simply does not report the unpleasant stimuli. It would appear then, that both perception and preference are functions in which considerable conscious control can be exercised, and the individual need only consciously suppress rather than repress to alleviate the anxiety associated with threat.

The results in the area of preference give further evidence for the relationship between preference and personality. But even more, it suggests that personality dynamics may be ascertained from a knowledge of the individual's preferences. Functioning as it does at a conscious level, this technique might well be used as an adjunct to projective methods that elicit unconscious responses. Utilized in this manner, it would allow for the differentiation between conscious and unconscious traits. The preference test might also be used as a measure of therapeutic change involving the

awareness and acceptance of previously repressed tendencies. The method might easily be broadened by replacing the hostile symbols, so that other traits such as dependency or homosexuality might be measured.

Because of the many investigations that have demonstrated the relationship between repression and recall, it was expected that here, even more than in the other areas of psychic functioning, Eriksen's hypothesized dichotomy would be found. In contrast to this, significance was not even approached. It is of interest to note, however, that only seven of the sixty experimental Ss recalled more hostile than non-hostile symbols. This might be attributed to a greater learnability or familiarity of the non-hostile symbols, but it seems more likely that this is due to a generalized inhibitory reaction in our society. Highly emotional moral attitudes center around the expression of hostility, so that individuals feel especially guilty of any infraction in this area. Since our culture allows for few direct outlets, it would appear that a generalized repressive reaction has resulted.

The third area of study, projection, brought out results that were inconsistent with the findings of Eriksen. In contrast to his theory, a perceptual defense-like reaction did not occur and the Rorschach was able to assess both conscious and unconscious hostility. Thus, the Rorschach uncovers the needs of the S, and they will be manifested regardless of the individual's defensive structure. This is in agreement with the research of Walker (1951), who found that even the repression of hostility did not preclude the perception of hostile stimuli on the Rorschach. However, a conscious element is also present, resulting in the hostile insightful Ss giving more hostile responses than the repressed Ss. Since the difference is not significant, it is necessary to follow the technique suggested by Klopfer et. al. (1954). The examiner must question those conscious attitudes

that are suggested by content analysis in order to distinguish between conscious and unconscious traits.

It should be emphasized that the experimental Rorschach method used was quite different from typical Rorschach procedure. It is believed, however, that generalization from this study can be made. Despite this difference in technique, the results obtained from each S suggest that this could be a useful procedure that might be applied to low responders while testing the limits. Not only would this allow for the accumulation of additional material, but it could also force deviant responses and make for an effective means of diagnosing troublesome borderline cases.

Since Eriksen's constancy hypothesis was not verified by this study, the results suggest that not only the individual's defensive structure must be considered when predicting behavior but also the environmental situation. The response of the S is not constant but dependent on the stimulus evoking the response, the type of response required, and the situation in which the response is elicited.

Concerning the stimulus, it would appear that much of the response is dependent on the degree of ambiguity of this stimulus. If it is obviously hostile, such as a gun, the S can, at a conscious level, accept or reject this. When the stimulus is ambiguous the meaning is unclear and the individual will be unable to deal with it at a conscious level. The need will then be manifested regardless of whether acceptable or repressed. Consequently, in this study, the defenders were able to reject the hostile symbols on the preference test, but on the Rorschach even the unacceptable needs were exposed. This would explain why some investigations have found differences between the Rorschach and more structured tests, such as the TAT (Hafner & Kaplan, 1960). It bears restating, however, that the differences will occur only with the defenders and not with sensitizers.

The type of response required or the psychic function involved also determines the reaction of the individual. Certain functions, such as preference and perception appear to operate at a more conscious level than projection. If the stimulus alone were creating this reaction, it would be expected that those rather obvious hostile responses often elicited by the Rorschach would be amenable to conscious control and thus suppressed. But unlike the reaction to the preference test these responses were not inhibited.

It would appear, then, that the more conscious the function the more the insightful Ss will respond to threat with an increased awareness and response to the threatening stimuli. The repressors, however, will react with less awareness and response, but when the function is less conscious, all Ss respond with and expression of their needs.

SUMMARY

The purpose of this study was to test Eriksen's sensitization-defense hypotheses, namely, that individuals respond to threatening stimuli by either an increased or decreased awareness and response to the threat. He further indicated that this reaction was constant regardless of the response elicited or the stimulus situation involved. In this investigation the anxiety provoking trait studied was hostility, and the psychic functions involved were recall, preference, and projection.

Three groups of 20 Ss each were selected from an original sample of 250 by means of a partial paired comparison technique. Through self and peer ratings the following three experimental groups were chosen: Ss rated hostile and aware of it; Ss rated hostile and unaware of it; and Ss rated average in hostility who rated themselves at this level. These groups were then tested and compared on tests of recall, preference and projection.

The following results were attained:

1. The preference test was the only measure to demonstrate Eriksen's sensitization-defense dichotomy. The hostile insightful group preferred significantly more hostile symbols than the hostile repressed group.

2. The test of recall was not effective in differentiating the hostile from the non-hostile group or the insightful from the repressed.

3. Content analysis on the Rorschach demonstrated statistically significant differences between the hostile and non-hostile groups. In contrast to Eriksen's hypothesis, the hostile-repressed Ss projected more hostile responses than the non-hostile group. This would suggest that the needs of the S are projected onto the Rorschach regardless of the individual's defensive structure.

4. Individuals did not show constancy of response, that is, they varied their response, either sensitization or repression according to the stimulus situation and the response required.

5. This study demonstrated that Eriksen's sensitization-defense dichotomy exists only in some circumstances, and the defensive method used is not constant but varies with the stimulus situation.

REFERENCES

- Eriksen, C. W. Implications for TAT interpretation. *J. Pers.*, 1951, 19, 282-288.
- Eriksen, C. W. Defensive forgetting. *J. Exp. Psychol.*, 1952, 44, 442-446. (a)
- Eriksen, C. W. Defense against ego threat in memory and perception. *J. abnorm. soc. Psychol.*, 1952, 47, 230-235. (b)
- Eriksen, C. W. & Browne, C. T. An experimental and theoretical analysis of perceptual defense. *J. abnorm. soc. Psychol.*, 1956, 52, 224-230.
- Eriksen, C. W. & Lazarus, R. Perceptual defense and projective tests. *J. abnorm. soc. Psychol.*, 1952, 47, 302-308.
- Hafner, J. A., & Kaplan, A. M. Hostility content analysis of the Rorschach and TAT. *J. proj. Tech.*, 1960, 24, 137-143.
- Klopfer, B., Ainsworth, M., Klopfer, W., and Holt, R. *Developments in the Rorschach Technique*. Yonkers: World, 1954.
- Lazarus, R. S., Eriksen, C. W. & Fonda, C. P. Personality dynamics and auditory perceptual recognition. *J. Pers.*, 1951, 19, 471-483.
- Mainord, W. A. Experimental repression related to coping and avoidance behavior in the recall and relearning of nonsense syllables. *Dissertation*, 1956, 1719.
- Phillips, L. and Smith, J. G. *Rorschach Interpretation: Advanced technique*. New York: Grune & Stratton, 1953.
- Rapaport, D., Gill, M. and Schafer, R. *Diagnostic psychological testing*. Chicago: Yearbook Publishers, 1945.
- Rosenstock, I. M. Perceptual aspects of repression. *J. abnorm. soc. Psychol.*, 1951, 46, 304-315.
- Rosenzweig, S. The experimental measurement of types of reactions to frustration. In H. A. Murray (Ed.), *Explorations in Personality*. New York: Oxford University Press, 1938.
- Schafer, R. *Psychoanalytic interpretation in Rorschach testing: Theory and application*. New York: Grune & Stratton, 1954.
- Walker, R. G. Manifest hostility and Rorschach and MAPS Test performance. *J. proj. Tech.*, 1951, 15, 444-460.

Received November 21, 1961

Projective Case Study of a True Hermaphrodite¹

EVELYN CRUMPTON

V.A. Neuropsychiatric Hospital, Los Angeles

The person who produced the following Rorschach and TAT material is a true hermaphrodite. No person is either wholly male or wholly female. The person called H is merely better balanced than most of us. He is presented to you not as the freak he is sure we will think him, but as a potential contributor to the projective understanding of sexual identification.

Born in the Midwest, of middle-class parents, H lived as a male during his early years, but his condition was known. As a boy he asked his father about himself and sex, and his father told him he was "just different". He had a small penis, obesity, large breasts, and only slight pubic hair. He left high school after his third year; as a 17-year old in the World War I Army, he was "teased but got along." Surgery for a hernia revealed part of the physical basis for his condition. After surgeons removed what was reported as "female sex organs", he became even more feminine, with large breasts, feminine body configuration, smaller penis, and labia-like scrotum. (He puzzled physicians as late as 1959, when the report read, "Evidence is overwhelming that ovarian tissue is present, but where?")

H describes his problem: "When I got out of the Army, I looked like a girl of 21. I tried to hide my breasts by putting on a large girdle that pulled them flat. My penis went back in my belly, and my scrotum almost disappeared. I had a high voice and no beard. A psychologist told me to live like a woman, so I did."

Since then H alternated living as man or woman, more often choosing

to be a man for the social advantages. He has lived in both Y.M. and Y.W.C.A.'s. He found more outlets for his intelligence and creativity as a man, working often quite successfully in occupations requiring both artistic sensitivity and technical know-how. In his mid-thirties he married a European princess who knew about his condition, but they were divorced after "four years of companionship", so that she could have a normal sex life.

At 50 his attempts to make a normal life for himself failed. "I was picked up for impersonating a man; if I went into a bar I was called a Lesbian." He tried to throw himself under a truck, and he was committed to a state hospital as actively psychotic, with delusions of persecution, ideas of reference, and severe guilt feelings.

Here the medical report stated, "Evidence of testicular failure, eunuroidism, manifested by atrophy of both testicles, the right one still up in the inguinal canal, almost absence of hair on face, and abnormally small amounts of hair on rest of body, with feminine type of pubic hair distribution, elongation of the lower extremities with lower span much greater than upper span."

After seven years of hospitalization, H made his way from state hospital to veterans' hospital. He made one more try at the outside world, but quickly drank poison and was re-hospitalized on a male ward as a paranoid schizophrenic. At the time of testing he was in his early sixties, functioning well inside the hospital. He could pass for either an effeminate man or a masculine woman. Bitter but not completely lacking in hope, H asked one question of the test material: "Am I nuts? Do I have a child-

¹From the Veterans Administration Neuropsychiatric Hospital, Los Angeles, and the Neuropsychiatric Institute, UCLA Medical Center.

ish mind? Or a mind that is equipped to go back out and battle the world again?" The testing itself upset his always unstable equilibrium, and he asked for a transfer to the women's ward.

Klopfer, Crumpton, and Grayson (1958a, 1958b) evaluated the commonly used projective indications of disturbances in sexual identification. The aim here is to use H as a test case for the validity of these clinical assumptions. H is known to have severe disturbance in his sexual identification, so severe that he is never quite sure what he is. His defective gene structure has affected his behavior as much as his biology, and we would expect him to project his disturbed identification onto the Rorschach and TAT.

To evaluate the Rorschach, we look for symbolically meaningful content, signs of defensive operations, breakdowns in reality-testing, inefficiencies in intellectual functioning, departures from usual modes of reaction. We want to distinguish signs of disturbed sexual identification from indications of disturbance in other areas, since we know that H is a depressed paranoid schizophrenic. We look specifically at responses with human content, responses with animal content, reactions to blot areas that frequently elicit human content, responses with specifically sexual content, and responses to blot areas that are most likely to elicit sexual content. Since his hermaphroditism has molded his life, we expect to find cues throughout the Rorschach, with at least some of these cues pointing to the most severe degree of disturbed sexual identification.

First the responses with human content. H has two, gossiping women on III and girl-cherubs on VII, both female. Seeing women on Card III is not unusual. H spontaneously identifies their sex; he dresses them appropriately; and he puts them to women's work. Then he calls them sloppy and makes sure we know he

has the vocabulary of a woman by referring to suspants. His ambivalence shows, but it is under good control. The intensity of his self-hatred breaks through when he follows the women response with a devouring, blood-sucking spider. It is as if he were saying, "It is the hostile female in me that makes me destroy myself."

We expect him to see females in Card VII, and he does, again identifying the sex spontaneously. The action of kissing each other is acceptable for cherubs. He does not see real adult women, a mild sign of disturbance. Once again the response which follows this human response seems to be related to his attitude toward women: upper dentures to Card VII makes it hard to avoid the association of vagina dentata.

The severest level of disturbance in sexual identification is indicated by seeing one figure as both male and female or as both human and animal, but this H does not do.

Since H gives only 2 human responses and 14 animal responses, we suspect he may be expressing some of his attitudes toward people at this safer distance. He sees a bat, crab, butterfly, butterfly tail, bear cub, spider, butterfly, monkey, fly, bat, cat, wolf, deer, and tarantula. Probably classification of these animals as symbolic of male and female would have low reliability, but the only clearcut aggressive male that seems to be present is the wolf, and it is immobilized in a crest. The Rhesus monkeys are looking down to see if they still had their tails. Along with the ragged bearskin and the seacow with unusual tail assembly, the monkey concept seems to reflect H's feeling of having been castrated. The three butterflies may represent the soft, passive side of women, and the spider and tarantula the dangerous, destructive aspects of femininity.

Possibly some of H's reactions to people may be seen in how he handles the areas which are often identified as people. On Card I he gives no re-

sponse or sign of disturbance to the center D which is often seen as a headless woman. Similarly on Card II he does not see people in the W, but there is no sign of disturbance. It is conceivable that the sometimes monster on Card IV may have prompted the paranoid "fly looking at me." He uses the bottom D area on Card IX, sometimes seen as a man's head, only as the unspecified base of a design; and he sees a shoreline in the pink D on Card X, which could be seen as a child. If we knew only his reactions to these cards, we would certainly not suspect any disturbed sexual identification.

H gives 2 responses with some sexual content, breasts on the cherubs of Card VII, and a female pelvis on Card X. The breasts are used merely to identify the sex. The pelvis is in a dr area and is explained by a defensive reference to personal experience, which suggests disturbance.

H gives no specifically sexual content to any of the "usual sex areas": center d and bottom D on II; "penis" and "breast" projections on the usual figures on III; top d, side d, and bottom D on IV; top D, center area and bottom opening on VI; center d on VII. He shows the greatest signs of disturbance to the usual "penis" area on Card VI. He comments, "I can't figure out what that doodad at the top looks like," and makes it something the bear was eating when it was skinned, which is certainly a strained connection. Later he makes the area a totem pole and unusual tail assembly of a seacow, expressing in mild form his over-evaluation of the penis and feeling of having been castrated. He responds immediately to the bottom red D on Card II, calling it a tail of a butterfly and denying the color. The "penis" area is identified as a skirt fluttering out, which is appropriate to the sex of the figures seen. On Card IV he cuts off the usual sex areas, but the disturbance he shows seems to involve the entire card rather than just these areas. On Card VII he does not

refer specifically to the bottom d "vagina" area when he calls the blot upper dentures. His responses are symbolically appropriate to the blot areas.

The most noticeable characteristic of his TAT stories is his frequent use of the first person. Certainly he is exhibitionistic. He has been looked at so long that he takes pleasure as well as pain from the looking. But also by saying "I" he effectively avoids the problem of identifying the sex of his character. He leaves us in the same doubt that he is in. In his other stories he does not misidentify the sex of any of the figures, or show any particular concern about what the sex might be. His identification of the hands of 18BM as belonging to the man's wife is unusual but plausible in his story.

He reveals his shifting identification most clearly when he unmistakably identifies with the woman of Card 4, saying "What have I done to make you so unhappy that you want to leave, another woman?"—and with the man of 18M who wonders, "Am I not a man?"

In his other stories containing both male and female, his involvement is less extreme, but he refers to the woman first on Card 2 and Card 6BM, and the man first on Card 18BM.

He expresses in several of the stories what seem to be his ambivalent feelings about himself. On 17BM he tells the story of a crippled boy who overcomes his handicap and climbs the rope, but on Card 3BM he says, "I feel that right now my life is at low ebb." When he speaks on 7BM of "the sorrow of a father for his son," he seems to be referring directly to his own father telling him that he was "just different."

H's Rorschach contains many signs of his disturbed sexual identification. It could be the record of a man with marked feminine orientation, intense ambivalence toward women, and castration fears. Or it could be the rec-

ord of a woman who identifies with her own sex, who envies men but considers them weaklings compared to the dominant, aggressive female. What the Rorschach of H does not show is the severity of the disturbance, as determined by the usually expected signs: symbolic or overt misidentification of the usual sex areas, combinations of male-female or human-animal in the same concept, or signs of gross subjective or objective disturbance to sex areas or references.

One might predict hermaphroditism or at least disturbed identification from the content of some of the TAT stories, but even here we do not see the severity of disturbance that might be expected. In both Rorschach and TAT he misses good opportunities for expressing his predicament. We cannot say that H has learned to live with his hermaphroditism and therefore to reduce the severity of the conflict projected onto the tests. If this were true, we would not find the paranoid and depressive signs that

pervade the record, and we would not find H in an NP hospital.

SUMMARY

Rorschach and TAT protocols are presented in a projective case study of a true hermaphrodite, whose see-saw solutions to the problem of sexual identification led him both to success as a creative professional person and prolonged hospitalization as a paranoid schizophrenic. Some popular signs of disturbances in sexual identification occur in the test material; many do not. The aim of the case study is to further the projective understanding of sexual identification.

REFERENCES

- Klopfer, B., Crumpton, Evelyn & Grayson, H. M. 1958a. Rating scales for ego functioning applicable to diagnostic testing. *J. proj. Tech.*, 22, 70-81.
Klopfer, B., Crumpton, Evelyn, & Grayson, H. M. 1958b. Manual for rating scales for ego functioning applicable to diagnostic testing. Los Angeles: UCLA Students' Store, (Mimeo., 75pp.)

Received April 17, 1962.

T.A.T.

1.

At last it's mine. I've wanted one for a long, long time. Father decided at last to give me music lessons. And now I've got my violin. How long will it take me, I wonder, to learn how to play this fine instrument. I shall practice hours at a time. At least an hour each day after school, until I become efficient. (If you were going to give a title to this story, what might you call it?) I'd say "An Answer to a Dream".

2.

(Very long reaction time.) This is a hard one for me. Well, let's see how would I word a story. I would say that the title of this particular picture would be "Will Education Take Me Away from the Farm." From the looks of the picture I'd say that the lady's resting after a day's labor. The child returning home from school and the father planting, working the — working the horse in the field — Evening time will come, work in the fields for the day will be done. Back home, washing for supper. I'm stuck. That shouldn't stick me but it does.

3BM

I'd say the title of this picture would be "The Drug Addict". It's a long trail back, the withdrawal time. Will I be able to make it, or won't I be able to make it. This is the hardest part. Drug addiction. I feel that right now my life is at low ebb. The pains and mental tortures and everything that goes with it, have got me in this particular position. I don't ever think that I will ever try drug addiction again. At last I've broken the habit. I'll never return to it again. I've never been a drug addict but I've read of it. (What about an outcome for this story?) This one? You mean ultimate results? (Just how things would work out?) At last I've broken the habit.

4.

Well, you could give this a dozen different titles. (Lp) Gollee, you could make ten different stories out of that. The title of this one is "Must You Go". What have I done—to make you so unhappy that you want to leave. Another woman? Tell me it isn't so. Actually you could write more than that. Short story, novel length. This all wrapped

up in that one picture. So many things in this picture—oh—She's pleading and he wants to get away. Perhaps he was pleading and she wanted to get away. And in order to try and save this marriage she began to plead all over again—because he'll be a hardworking man. He's not a drinker. He doesn't have a facial picture of an alcoholic. He's a clean cut, God fearing man. She's a good woman and a good wife but perhaps small differences present arguments. This is spur-of-the-moment action without thinking. (Looks at watch.) Yet in another they'll be a very happy couple because he believes in her and she believes in him. I'm a horrible story teller on a spur of the moment.

6 BM

On this one I believe about a dozen stories could be made for this particular picture. One is that her son might be leaving for the Army and she just doesn't want to shed a tear right now, not while he's here; and again perhaps he's telling her some frightening news—about an automobile accident. Perhaps he's done something to make her very unhappy. Perhaps she has told him something—very bad news. Perhaps his father just passed away. She's trying to tell him. (Vlp) Oh me. (Vlp) Oh. (Vlp) I don't think I am getting any more out of that.

7 BM

In this picture I'd suggest a title "Father's Advice to His Son". The father's explanation of life to his son. (Vlp) I don't know whether I see tears in that man's eyes or not. It does look a little bit like there's a tear in that man's eye. The sorrow of a father for his son. A sorrowing father and his wayward son. It could be a happy father at his son's graduation.

17 BM

Whoops! The title of this picture would be "I Could Do It". Perhaps this boy is crippled and wants to show how he can climb up the rope, hand-over-hand without the use of his feet. There's a look of exaltation on the boy's face. This is perhaps one of the greatest things he's done before in his life. This could be a crippled boy's physi-

cal prowess with a look of success in his eyes. I'm a horrible story teller.

13 BM

This presents a problem. You could make a half a dozen stories out of this. On this picture I'd say that the title of it "What Have I Done". This picture could denote betrayal. It also could denote that the man refuses to betray the chastity of the woman. Another title could be "Am I Not a Man". Is that enough on that one I hope?

8 BM

The title of this picture would be "The Dreamer". Dreams of great things, of becoming a great surgeon through education and knowledge. Through knowledge to do. Today I've made up my mind. I will be a doctor and a great surgeon.

12 M

I'd say the title of this picture would be "Good-bye Friend". There are a half dozen good titles to this particular picture. On this one there's a—could be a long story about two boys who grew up together and ah—when there came a part of—time of their lives that they were separated or to be separated one would be bidding the other good-bye. One of the boys didn't want to let the other know what he was going to be so that the other boy might attain success. Wonderful, intriguing.

18 BM

This picture would be about two or three titles. I'd say that this picture represents a man who was alcoholic, probably drunk, and his wife was trying to keep him from going out and spending his money foolishly in some bar room. She is grasping him by the shoulder and also by the arm to prevent him from going out again. Without a doubt the man is already pla—ah drunk.

(Patient was asked to choose one card from 10, 13B, 14, 20.)

10

This picture represents love. This title of this picture should be "I Love You With All My Heart".

RORSCHACH

I

13"

1. This one 11 a bat in flight. (Inverted card, then turned upright.) (Vlp) To me it 11 a bat in flight. (W)
2. At the same time it also 11 a crab with

its mandibles sticking out. (Blot-description and comment identifying occupation was deleted.) (Pt. laughs.) (W inverted)

(This you said 11 a bat in flight.) You want me to take a look? (Yes.) Well, it does 11 a bat in flight, but it also

can be reversed, it ll a crab.

Because there are normal crab mandibles there (upper center projection) and on a crab the eyes are protruding (top center, small projections pointing up). (I'm getting behind, let me catch up.) Oh, I'm sorry. Now. Have you got that down where the eyes and the sockets up there protruding? (What makes it look as if they are protruding?) Well, sometimes in crustaceans, the eyes are set on knobs. I don't know the biological term for it. And yet

Add 1: Perhaps it is defined as a butterfly also. (W) (Tell me more about it, what about it ll a butterfly?) Well, if you'll notice, some butterflies have markings such as these on their wings. (Such as what?) These. (Points to extreme outer edges.) And down here ll the trailing edge on the butterfly's wing. (Inverted card.) (Is there anything else about the butterfly?) Well, actually it is an old butterfly. (An old butterfly?) Yes, one that is ready to die, for the simple reason that its wings are all broken. It's the end of the season. (When you say the wings are all broken, where are you referring to?) Right here, its broken edges. (Points to outer edges and inner holes.)

(When you said about the bat in flight, what made it ll a bat in particular?) This part here has a bat's wing structure more or less. Ll it had its wings outstretched. (Anything else gives it a look of a bat?) Well, possibly the body. (Lp) I wouldn't know what these fragments are (tiny dots outside of blot) around here or what to define them as.

II

16"

1. Well, it has the tail of a butterfly. (Bottom red)
2. Ll two bears, bear cubs playing pitty pat. Sitting on their haunches. (Inverted card, then turned upright.) How dumb am I? (Doing this makes you think you're dumb?) Very dumb. (W)

(The first thing you mentioned here is the tail of a butterfly.) That's this one. Well let's see. I still see only the little cub bears sitting on their haunches playing pitty pat. I don't

know what else to call it. (Laughs) As to the butterfly, I don't know what kind of a butterfly it is because I've never studied butterflies. And it ll the bears are looking up at some meat. (Upper red) Each one of them has got their noses up there, and little faces up there, and looking at something that might be considered meat. And they stop in their game of pitty pat to stare at the handout someone is giving them. (When you refer to the meat there, what is it that gives it the look of meat?) Oh, these stripes here. (Anything else?) And it ll there's some fat in there, too, the light spots. But I don't think bears bother with fat and lean, they just like the meat, period. (About the bears, any particular thing about the looks of the ink blot that makes it ll bears rather than say any other kind of animal?) Well, no, could be rabbits but it's not. Because of the leg structure there and the leg structure here are seemingly the same. Bears will play like that and guinea pigs won't (I) (Another question about the butterfly, you mentioned just the tail?) Well, this is the wing structure here too. It runs up into the inkblot. The body structure. What appears to be the tail of the butterfly but actually is a part of the wing structure. (Is there anything else that suggests butterfly or does that about do it?) That about does it.

III

1. Well, this ll two women gossiping over a pot of stew. (Inverted card then turned upright.) (W except red)
2. When upended it looks just like a spider was staring you right in the face. (Pause) With it's two front legs upended and its mandibles. Ll it's about ready to devour (W)
3. A butterfly (center red).

(You mentioned several things here, first the two women gossiping over a pot of stew.) Well it could be a pot of stew or could be their washing or something. (Could you describe the women a little more?) Well, they are eggheads (laughs). They've got protruding noses and mouths. They have elongated necks, and their breast formation is full. And the jackets they're wearing are full sleeved or three-quarter sleeved. And one is

carrying laundry or about to put something in the pot. In fact they both have identical things to place in the pot or hang on it. Or they could be carrying their purses. And each one is wearing a skirt two inches above the knee and fluttering out. Both are wearing stockings. Seems to be very sloppy, a garter must have busted on their suspants. Each is wearing high heeled shoes, spikes. Both of the jackets have pockets in them.

(Upended it ll a spider?) Ll a tarantula. With the eyes staring at you, the mandibles open and about ready to devour a butterfly that's probably lost its way and gone into the web. (What suggests butterfly?) (Points to center red.) (Anything about the looks of it?) (No response.) (What I mean is, could you describe it more?) Well, it has a very small body and two extra large wings about four times, about three times the size of the body itself. Spiders don't eat like we eat. They suck the blood out of the victim and leave the rest, but in this case it ll the spider's already had a bite out of it. And I wouldn't know what these other two blots might mean except it might have been previous meals.

Add 1: Actually those spots ll a couple of little Rhesus monkeys looking down to see if they still had their tails (laughs). (Outer red)

IV

Oosh!
15"

1. This ll a bear skin rug. (W except outer side projections)
2. And then again it ll a fly looking at me. (W except for usual "boot" areas inverted)
3. And upended it ll the crest of my wife's stationery. She was a Hapsburg. (Card inverted) (Turn card upright) Glory be. I never thought I'd come to this (laughs). (You mean taking these inkblots?) I used to see something different in them every time I saw them. (W)

(This bear skin rug was the first.) (Patient turned card in every direction.) Well now, the top of the head, the back of the head ll the view you'd normally look at a bear. Bear skin rug. The ears sticking out here, a tuft of hair here. Ll the bear skin rug goes up about like so, about 1/3. I don't

know what you'd call this (points to outer side projections). (When you said about one-third, did you mean just the top?) You see the bear skin rug goes clear down to here. Whoever has done the skinning has made a slight error, didn't cut it right. Seems to have been skinned by an amateur, who did a very ragged job in the whole thing. And yet when you look at it again, it ll the head of a fly. These would be the two eyes staring at you. (Where?) These. (Points to outer side projections.) (Back to the bear skin rug again, is there anything else about it that gives it the look of a bear skin rug?) Yes, you have the shading in here that could be the look of skin perhaps. (On the fly, is that the whole thing or part?) Only part. Actually this would be the proboscis down here (center bottom). (Here?). Well, I wouldn't know what to call that in here (usual boot area), that still ll part of the bear skin. But this resembles a fly in here, and here and here (pointing to remainder of blot).

(On this one you also had the crest?) (Inverted card.) This ll a crest this way. (What about it?) On the Hapsburg crest they have that old crow in there. (Laughs) I don't know what else to call it. That's what I use to call it on my wife's stationery. An eagle with some sort of crown on the head of the eagle. And of course, if that were on a shield, it would look just like a crest. (Anything else about that one?) No.

V

2"

1. This one really ll a bat. The antenna, the wings, the tail or feet, wings outstretched. And upended it looks the same like a bat was hanging from a rafters. (W)

(The bat?) The bat. I see nothing in it but a bat. One way or the other. (Is this the same bat or two different bats?) One way or the other it still ll a bat. (Anything else that gave it the look of a bat besides what you mentioned before?) No.

VI

8"

1. Well, this ll a bear skin rug again. (Inverted card then turned upright.) I can't

figure out what that doodad at the top is. Anyway it has whiskers, whatever it is. Maybe it's something the bear is eating when it was skinned. (Comment identifying occupation was deleted.) (W)

2. Well this one could have been cut from a seacow or something. Might have had an unusual tail assembly. (W)

(This is a bear skin rug?) A bear skin rug with

(Add 1:) an Indian totem pole (top D) on the end of it. (Could you describe this one more?) Well this gidget here—could be some part of a pelt but what part? (Pause) I wouldn't know what to call it unless it would be some type of a tail section. (What about the rest of it, what makes it is a bear skin rug?) Well, it has the shape, the coloring, the depth of black and greys, it has the coloring along here. (On the other one that had a bear skin rug, you mentioned the shading, is that what you mean by coloring here or do you mean something else?) Yes, the shadings. You see there are 117 shades of black to white. Artists call it degrees of color and (Blot-descriptive comment identifying occupation was deleted.)

(You mentioned a totem pole?) Well, this would be this part here. (Describe it more, what makes it is a totem pole?) Well, it's tall and it has outstretched wings of a thunderbird and, of course, people up north generally with totem poles ordinarily put the entire family history on a totem pole and this has a degree of totem pole look, the outstretched wings, and that would be the ancestry of the clan.

(You also mentioned something cut from a seacow?) Could have been. A seacow has an unusual tail structure, and I thought that this might have been, the unusual tail structure.

VII

Oosh! 18"

1. Well, this is a couple of cherubs about ready to kiss each other. One's got her face all puckered up and the other one is ready to kiss. (Upper 2/3)
2. Upended it is some upper dentures (smiles). (W inverted)

(A couple of cherubs about to kiss each other?) Yeah (laughs). It really

does it? (Could you describe them more?) Well they have the facial characteristics of a cherub. They also have where the eye and the nose comes around there in cherub fashion and the jaw structure. This one is waiting to be kissed (on the left) and this one got all puckered up ready to jump over and kiss. (You mentioned just the heads. Is that all or is there more?) Here would be the body with the arms outstretched. (Anything else?) Yep. The shading there is similar to that of the breast, the shading is the same (pointing to the right figure). So they both must be girls. Add 1: These down here are cats. (Bottom 1/3) One's facing this way and the other is facing the other way. (Anything to do with the cherubs or is this something different?) Just they're watching the cherubs above. (What makes them are cats especially rather than just any kind of animal?) This kind of shape here, the eyes, back, the head, the fur. You know how a cat's laid out on the floor, tucking their tails up under them. You can see markings here, the shades. (The other thing you mentioned was the dentures.) And the dentures. When you turn it upside down. (What gives it the look of dentures?) Oh, the shape of it. (Points vaguely at the whole inkblot.) I don't know, the upper part of it. (Anything else or does that about do it?) That does it.

VIII

Hm, this is done in art work. Well let's see, this is done in magenta, chrome yellow, and cyan blue. Uh, gee, I don't know, well.

20"

1. Over here we have wolves. (Side pink D.)
2. This is a crest (remainder of W). Wolves on this side. An insignia down here, background, flags, a crest, of a crest. (Lp) This looks to be like a family crest, a crest of a country or something, because it has the sole, the plate, the insignia, the flags, and the top part of it, whatever it would be.

(You described this pretty much before. About the wolves part of it, what about it makes it look especially like wolves?) The shape of the animals. Of course it could be a bear, but the legs are too long for a bear.

11—I would surmise it to be a wolf. And then the plate down here is an ensign or insignia. Two flags (blue area) and the top part, of course, would be the top part of a crest. I don't quite understand what this would 11 (orange). (Which?) Down here in the magenta. This 11 the folds of the ensign. (The folds?) Well, you see in dress material, how it folds in and out. (You mentioned the ensign in the magenta, did you mean the color of it or were you just showing me where you meant?) Well, it's made up of three colors. It's made up of magenta, cyan blue, and chrome yellow. (Blot-descriptive comment identifying occupation was deleted.) The flags are represented as being in motion by the folds. (Add 1) The top part, the top of the crest could be a mountain, a snow capped mountain. (Top D) (What makes it look snow capped?) But it's truly the crest. (Well, when you think of it as perhaps 11 a snow capped mountain, I wondered what it was that made it look snow capped especially?) Well, it would be the light areas in here. That's all I see in that.

IX

22"

1. Well, I see two deer. (Orange D)
2. I see marshland. (Green D)
3. I see a torch, a flame against a blue sky. (Center D, S)
4. But I don't know what the magenta 11. I can't place this, unless it represents the base of the design itself. This 11 some marshland and these 11 deer. I can't place this. (W)

(First you mentioned the deer?) The deer with antlers. (Where are the deer?) Right there. (Describe them a little more.) They seem to have, as described here, would be classed as jumping. And the territory or green part down here would be either territory or it could represent America or it could represent Canada or it could represent Alaska. And here is a torch here with a flame. (What gives the impression of flame?) Well, actually in this case the flame would be—it happens to be here a definite green in there (Blot-descriptive comment identifying occupation was deleted.) and the chrome yellow on top of that

which made it a green flame against a blue background, and I would say that would be a cerulean blue (Blot-descriptive comment identifying occupation was deleted.) And, of course, if this were an area of country, this would tend to be either wooded or mountainous because of the dark shading. (Usual deer head area inside green.) And the lower part of this, I don't know what that would be unless it could be a stand on which this fits. Oh, I don't see anything else in there. Nothing else seems to make sense. (Turns card to every direction before putting it down.)

X

Oh, here's a pretty one. (Inverts card, then turns upright.)
31"

1. Well, (Reference to occupation was deleted.) this 11 the skeleton structure of the pelvis. I really wouldn't know what this is. (dr, center blue, part of adjoining pink)
2. This 11 shoreline. (Pink D)
3. This 11 flowers. (Outer yellow D)
4. This 11 spiders or what you call them, tarantulas, with something in their mandibles. (Outer blue D)
5. This top part 11 a clevis rod of some sort, fastened on a doohickey (Top grey D)
6. I don't have any idea what this or this is (outer brown, and bottom green) unless it would be string beans. (Bottom green D)
7. And this 11 that little gadget that comes off of one of these trees, that comes down, seed pod. I forget what tree it grows on. That's all I can see in it. Do I have an imbecile rating now? (Inner orange D)

(You mentioned first the skeleton of a pelvis?) Pelvis, yes. Let's see. Hm (lp). (Where was the pelvis?) Right around here. (Points to center blue and the adjoining pink, the edges only going up.) (What gives it the appearance of a pelvis?) Well, I think the situation of these two images here (blue) tends to give it the look of a pelvis as you look down at it. Could be a female pelvis because there seems to be an arch here with little spots showing the method with which it opens. (And the shoreline?) Right here. (Where?) Right in the center of arch. (What suggest shoreline?) Oh, the shoreline would be right along here, the light

and dark areas of the edge of the magenta color. And these 11 two tarantulas, one on each side, or could be one of these big beetles, what do you call them, snapping beetles with mandibles stretched out or could be spiders or tarantulas. (You mentioned the light and dark areas on the edge, did that have any special significance?) No, other than it 11 a map or a topog of the shoreline. And this part 11 flowers. (How?) The chrome yellow there seems to be a blossom of some sort. And this part 11 string beans, the green. (When you say green do you mean the string beans are green or showing which area you're using?) Yeah, the string beans are green. (At the top you mentioned a clevis rod, what is a clevis?) Sort of like a wish bone (demonstrates shape). (You also mentioned a seed pod?) Here's the seed pod right here. I don't know what kind of tree it would come off of. (You mentioned before the type that comes down?) Yeah, that whirls down, you know (demonstrates twirling motion with hands). I don't know what these are for here, I don't think I was able to describe those (points to outer brown).

LIMITS

(Patient was asked if he saw any sex organs before that he had been reluctant to mention.) No. (Look now.) Uh uh. (Patient did look at card.)

(FM on card VIII) Well, they're just kind of standing there actually. Actually it's a part of a crest, because the way the animals are standing in the crest, they wouldn't be standing in real life.

(Card liked best) Oh—I think this one here (card VIII) is the one I prefer. (Why is that? It really represents—it really represents something. It's a complete picture actually. (How about the one you liked the least?) Oh—I don't really know other than I imagine this splotch right here would be the most unlikely, the most unintelligible (card I).

(Now I'd like you to do one more thing with the cards. Take them and sort them into two groups using any basis you chose.) (The patient sorts quickly) (What is the basis for the sorting?) This is something. I think. This group, these really represent something. (And you mean these others don't?) That's right. (Reference to occupation was deleted.) (Card VI, IV, and I in group not meaning anything.)

Interpersonal Measurement of Two Occupational Interest Groups

CHARLOTTE DAVID¹
University of Portland

Many theories have been advanced to explain why people choose the work they do and what they are working for beyond the need to earn money for the gratification of basic physiological needs. In recent years considerable attention has been directed toward the role of personality factors in vocational choice and satisfaction.

The relationship between occupational preference and different aspects of personality has perhaps been most thoroughly explored with interest groups. Tyler (1945), in studying the interests of college sophomores and their scores on the Minnesota Personality Scale, noted that students who have a tendency to avoid people tend also to have technical interests. Berdie (1944) related the number of likes and dislikes on the Strong Vocational Interest Blank to performance on the MMPI and found that people with an interest in occupations involving personal relationships were usually emotionally "acceptant" of their surroundings, while those who were interested in the "realities" of life tended to be somewhat more cynical. Teevan (1954) attempted to correlate personality factors with college curricular majors using the Blacky Pictures as a measure of personality. He noted that students in the social science division had higher disturbance scores on Oral Sadism, Guilt Feelings and Oedipal Intensity than did art, language, and physical science majors, while the students majoring in science had the lowest disturbance scores in almost every category.

Among the most active workers in the field of occupational psychology has been Anne Roe. In an effort to

find some of the personality factors which seemed to be most relevant to occupational choice she embarked upon a series of studies of research scientists in both physical and social sciences (1946, 1949, 1950, 1952, 1953). Out of her own work and the general body of information available in occupational psychology, she developed both a method of classifying occupations and a theory of occupational choice (1956, 1957). This classificatory system is a two dimensional one utilizing focus of interest as one dimension and level of responsibility and training as the other. "Classification by focus is clearly related to classification of interests" (Roe, 1956, p. 147).

Roe hypothesized that such factors as socio-economic background, intelligence and education are more clearly related to level than to focus, and that the personality factors are more closely allied to focus than to level. She found, for example, that physical scientists were more independent, less aggressive and rebellious, and also less interested in interpersonal relationships than are social scientists.

Several difficulties arise in attempting to ascertain precisely what role personality factors play in vocational choice. The number of different methods of gathering and interpreting the data make replication and the comparison of different studies difficult. There has also been a wide disparity in terms of the area of personality focused upon and the frames of reference used. Nowhere is there a really clear, well-rounded personality description of occupational groups which presents data in behavioral terms and which differentiates overt from covert behavior. For example, when Roe suggests that physical scientists are less aggressive and rebellious than social scientists is she speak-

¹This is an abstract of a doctoral dissertation submitted to the University of Portland and carried on under the direction of Sheridan P. McCabe, Ph.D.

ing in terms of underlying feelings, overt behavior, values, or self concept?

Statement of the Problem

It is here proposed that some further clarification might result if occupational groups could be studied within a single frame of reference which would allow for differentiation among various behavioral levels, and which would, at the same time, use largely objective easily replicable methods of scoring and interpreting the data.

The specific aim of this study was to attempt to determine the characteristic patterns, if they exist, of two occupational interest groups, using the conceptual system and methodology laid down by Leary in his book, *The Interpersonal Diagnosis of Personality* (1956).

On the basis of theory and the findings in previous studies the following hypotheses were tested:

1. A given occupational interest group will yield a characteristic clustering of scores at each Level on the Interpersonal Circle.
2. Two different occupational interest groups will yield significantly different clusters at each Level.

METHOD

Description of Tests and Scoring Methods

Leary's model of personality is a multidimensional one which presents a systematic method for viewing the interpersonal aspects of human behavior. He classifies this behavior in terms of a two dimensional circle, the horizontal axis of which represents Hostility-Affection, and the vertical axis Dominance-Submission. The circle is divided into eight segments representing characteristic ways of relating to others. These octants are further divided into adaptive and maladaptive intensities. Thus, octant #1 is designated Managerial-Autocratic, the former being the adaptive description and the latter the maladaptive. Following are the octant designations with their numerical and letter codes:

Octant #1	AP	Managerial-Autocratic
Octant #2	BC	Competitive-Narcissistic
Octant #3	DE	Aggressive-Sadistic
Octant #4	FG	Rebellious-Distrustful
Octant #5	HI	Self Effacing-Masochistic
Octant #6	JK	Docile-Dependent
Octant #7	LM	Cooperative-Overconventional
Octant #8	NO	Responsible-Hypernormal

In addition to assessing types of behavior, this system allows for a multi-level approach to this assessment. Translated into behavioral terms, Level I is the way a person is seen by others, Level II is what he chooses to tell about himself, Level III is akin to preconscious or fantasy behavior, Level IV is that behavior which is so anxiety provoking as to necessitate its complete avoidance, and Level V is the level of Ego-Ideals. Operationally, Leary defines these levels in terms of the source of the data.

In this study Level I data were obtained from the MMPI. Since this test is a self description it is not a *direct* measure of behavior as seen by others. However, correlational studies (Leary & Coffey, 1954) showed a significant relationship between the prediction of behavior as measured by certain MMPI scales and actual interpersonal behavior as rated by others. The clinical scales alone were found to be ineffective as predictors of behavior in a normal population, but several non-clinical scales were fairly adequate. Thus, Leary (1960) suggested the inclusion of the Ego Strength (PgB) and Hysteria Denial (HyD) scales in the following formulae:

$$\text{Dominance} = \text{Ma} + \text{PgB} - \text{D} - \text{Pt}$$

$$\text{Affection} = \text{HyD} + \text{K} - \text{F} - \text{SC}$$

Data for Levels II and V were obtained from the fourth major form of

an adjective check list devised by La-Forge and Sucek (1955).

Level III data were derived from the following TAT cards: 1, 2, 3BM, 4, 6BM, 6CF, 7BM, 12M, 13MF, and 18BM. The TAT themes were rated according to numerical octant codes by a rater trained by Leary in the use of this scoring system. All measures used in this study were previously used by Leary.

Subjects

Two groups of male undergraduates were used in this study. One group, designated as Group PS, was made up of 72 students with standard scores of 45 or higher in Occupational Group Scale #II of the Strong Vocational Interest Inventory. The second group, henceforth designated Group SW, consisted of 80 students with like scores on Occupational Group Scale #V. In addition to meeting the above requirement, each S attained a score at or above the mean on the interest Maturity Level Scale of the Strong in order to control for the age factor and to insure homogeneity with respect to this scale.

Procedure and Handling of Data

Following the administration of the tests and processing of the Leary data, a summary point was computed for each S at each Level and plotted on the interpersonal grid. These summary

points were dealt with in terms of frequency of occurrence in each octant of the circle and treated statistically by use of the Chi Square Tests of Significance and *t* tests.

RESULTS

The following tables show the distribution of scores for both groups on the circular grid and will be used to show how the results are related to the hypotheses tested. Table I demonstrates that Group SW yielded significant clusters of summary points at Level I, II and III, but not at Level V. Thus, the hypothesis stating that this group would yield a characteristic clustering is supported on all Levels except the Level of Values.

Table II demonstrates that Group PS shows significant clustering at Level I only, suggesting that the above hypothesis does not hold up for Group PS as it did with Group SW.

Chi squares were computed to test the difference between groups. They were significant only at Level II, where *p* was found to be less than .01. Despite this fact, examination of Tables III and IV show that these groups do differ in terms of intensity at both Levels I and II.

Thus, the hypothesis concerning the differences between groups is supported at Levels I and II only.

The mean summary points for each group indicates that Group SW is

TABLE I—Distribution of Summary Points and Chi Squares for SW Group

Level	1	2	3	4	5	6	7	8	N ^a	X ^a	P
I	40	4	5	5	0	3	3	20	80	128.4	.001
II	34	16	3	0	6	3	2	16	80	92.6	.001
III	9	8	10	6	4	7	7	10	70	16.4	.05
V	7	10	12	10	9	14	9	9	80	3.2	NS

^a The smaller N at Level III results from unsuitability of some TAT stories for scoring.

TABLE II—Distribution of Summary Points and Chi Squares for PS Group

Level	1	2	3	4	5	6	7	8	N ^a	X ^a	P
I	27	7	7	9	4	4	3	11	72	46.88	.001
II	7	15	12	7	12	3	6	12	72	12.00	.10
III	9	7	11	14	4	4	7	4	72	12.53	.10
V	8	8	16	10	4	11	9	6	60	10.00	.10

^a The smaller N at Level III results from unsuitability of some TAT stories for scoring.

TABLE III. Means, Standard Deviations, and *t* for Love Scores

Level	Group PS		Group SW		<i>t</i>	P
	Mean	S.D.	Mean	S.D.		
I	54.19	12.66	67.72	8.21*	3.20	.01
II	51.55	9.99	57.85	7.50	0.34	.01
III	52.56	11.33	52.07	9.75	.20	NS
V	40.54	11.44	48.72	9.75	.10	NS

* *t* significant at .001 levelTABLE IV. Means, Standard Deviations, and *t* for Love Scores

Level	Group PS		Group SW		<i>t</i>	P
	Mean	S.D.	Mean	S.D.		
I	50.19	8.22	54.67	8.79	1.39	.01
II	48.87	9.90	52.12	7.00	2.24	.01
III	46.61	10.56	47.65	10.74	.51	NS
V	47.08	19.79	49.52	10.00*	.94	NS

* *t* significant at .01 level

classified as Autocratic at Level I, Managerial at Level II, Aggressive at Level III, and Self-Effacing at Level V. Group PS is seen as Managerial at Level I, Competitive at Level II, Aggressive at Level III, and Rebellious at Level V.

DISCUSSION

In discussing the results of this study it must be borne in mind that the norms used throughout were based upon a psychiatric clinic admission population because no others are presently available. It is probable that norms established on a college population would yield a more precise, clear-cut picture, though this cannot be established until such norms are available.

The experimental hypotheses, that a given occupational interest group would yield a characteristic personality pattern, and that the patterns of two different interest groups would be significantly different, are only partially supported. The multilevel pattern of Group SW is, except for Level V, a fairly clear cut one. If it were possible to extrapolate from the data a description of this group based on statistical summary points alone, the following picture might result: People with high interests in social welfare occupations tend to behave in ways which seem to others somewhat domineering and autocratic, but without hostility. They regard themselves

in much the same way as they are seen by others, but do not see themselves as being open as controlling as others think. Underlying this overt behavior is a measure of aggression and competitiveness which is more direct, within normal limits and therefore, presumably, under adequate external control. Their value system as measured by their ego ideals are quite diverse and seem more clearly related to individual differences than to any common group characteristics.

Any attempt to so extrapolate from the data of Group PS, however, is much riskier, unless the description is undertaken with tongue in cheek and limited to very general trends. Describing them in terms of Leary's system based upon summary points only, this group too, is likely to be seen by others as dominant. (This may be a characteristic of all college populations.) However, they are neither so controlling as the SW Group, nor as loving. They tend to see themselves as competitive and somewhat aggressive, while underlying their overt behavior is some rebelliousness. Again, there is no common ego ideal, but they tend to value skepticism. Such a description of Group PS must be viewed with a jaundiced eye, however. The amount of variability within this group suggests that it is a much more heterogeneous one than seems indicated by their common grouping

on the Strong. According to Roe's occupational classification (1956), the Strong Occupational Group #II encompasses at least two categories: Technology (IV) and Science (VI). In view of our findings it seems likely that her classification is the more adequate one for personality investigation.

In relating the results of this study to those of others, it is noted that Teevan (1954), using the Blacky Pictures, found that social science majors had relatively high disturbance scores on Oral Sadism, Oedipal Intensity, and Guilt Feelings. Since the Blacky is interpreted in terms of psychoanalytic theory it seems justifiable to compare these results with the Level III findings. The clustering of Group SW in the Aggressive octant at this Level tends to bear out Teevan's findings, though without direct reference to particular focus. On the other hand, assuming that Guilt Feeling would be reflected in Octant #5, the Self-Effacing-Masochistic category, inspection of these data does not bear out Teevan's findings at any Level. There are no measures in the present study to compare with Oedipal Intensity, though the question of parental identification or dis-identification could have been incorporated through the use of the adjective check list to describe father and mother. It would then have been possible to compare the relative positions of Self, TAT-Hero, Ideal Self, Mother and Father on the interpersonal circle. Incidentally, such a method could be used in studying the relationship between self-parent occupation and degree of identification with parents or parental figures.

In discussing the personality aspects of various occupational groups, Roe (1956) sets up a circular schema very similar to Leary's in many respects. In her model the vertical axis represents "Orientation-Toward-Persons vs. Non-Persons," and the horizontal axis represents Warmth-Coolness, as related to parental figures. On this circle she places the Service Group in the

quadrant delineated by Warmth and Orientation-Toward-Persons, the Technology Group in the Warm, Oriented-Toward-Non-Persons quadrant, and the Science Group in the sector defined as Cold and Oriented-Toward-Non-Persons.

It seems likely that the horizontal axis of both Roe's and Leary's models are much the same: Cold-Warm or Hostile-Affectionate. The relationship of the two vertical axes is, however, less clear. It is not certain that Orientation-Toward-Persons vs. Non-Persons is related to Leary's Dominance-Submission continuum; but if such is assumed then the results of this study at Levels I and II are in general accord with Roe's. That is, Group SW is more dominant, more affectionate, and less independent than Group PS. There is, however, no counterpart to the Level III findings of aggressiveness in the SW Group and a tendency toward rebellion and aggression in the PS Group. In Roe's studies of physical and social scientists (1953), however, she noted that physical scientists were less aggressive than social scientists. This relationship is found to be the case at Level III. It seems probable that the confusion which arises here is part of the general failure in personality evaluation to discriminate between feeling and action, and between the conscious and unconscious aspects of personality.

This study was undertaken with several ideas in mind. An attempt was made to find out which particular aspects of personality seem to predominate in separating out different interest groups, to relate these findings to those of previous studies, and to test the usefulness of Leary's concepts and methodology for this kind of study.

It is here suggested that Leary's model has real value for further investigations into the area of the "occupational personality". However, a word of caution is in order. Leary's findings (1957) on small groups of

undergraduates tend to confirm the general impression gleaned from this study that college populations cluster in the dominant half of Leary's circle, especially at Levels I and II. It is therefore suggested that more meaningful results would be obtained if norms could be established for college populations. A second suggestion relates to choice of Ss. It is felt that the meaning of the results is obscured and confused by the possible differences existing in the PS Group. Replication of this study should be undertaken with "purer" Ss than those used here. Inspection of the Strong's would seem to indicate that the interests of chemists and engineers is too disparate for true homogeneity and that clearer results would be obtained by using single interest scales rather than group scales.

SUMMARY

The aim of this study was to test two occupational interest groups using Leary's International System of Personality Diagnosis, with the idea of obtaining more information about the personality factors operating in vocational choice.

Seventy-two male undergraduates with standard scores of 45 or higher on the Strong Occupational Group Scale #II comprised the physical science group (PS), and 80 male undergraduates with like scores on the Strong Occupational Group Scale #V made up the social welfare interest group (SW). All Ss also had standard scores at or above the mean on the Strong Interest Maturity Level Sale. All Ss were given the MMPI, two interpersonal adjective check lists and the TAT. The data were then processed according to Leary's methodology and a summary point plotted on the Interpersonal Circle for each subject.

The following experimental hypotheses were tested:

1. A given occupational interest group will yield a characteristic clus-

tering of scores at each Level on the Interpersonal Circle.

2. Two different occupational interest groups will yield significantly different clusters at each Level. The Chi Square Test of Significance was used to test hypothesis #1 and Chi Square and *t* Tests were computed to test the second hypothesis.

Statistical significance in support of hypothesis #1 occurred at Levels I, II, and III in Group SW and only at Level I in Group PS. Hypothesis #2 was supported at Level II only, although *t* Tests showed a significant difference between groups at both Levels I and II in terms of intensity.

These results were discussed relative to their relationship to previous studies in this area. It was suggested that Leary's concept of the interpersonal measurement of personality had some merit for occupational psychology but that future studies be undertaken using different norms and purer groups.

REFERENCES

- Berdie, R. F. Factors related to vocational interest. *Psychol. Bull.*, 1944, 41, 137-157.
- LaForge, R. and Suczek, R. The interpersonal dimension of personality. III. An interpersonal check list. *J. Pers.*, 1955, 24, 94-112.
- Leary, T. and Coffey, H. The prediction of interpersonal behavior in group therapy. *Psychodrm. grp. psychother. Monogr.*, 1955, 28.
- Leary, T. *Multilevel measurement of interpersonal behavior: A manual for the use of the interpersonal system of personality*. Berkeley: Psychological Consultation Service, 1956.
- Leary, T. *Interpersonal diagnosis of personality*. New York: Ronald Press, 1957.
- Leary, T. Personal communication, 1960.
- Roe, Anne. A Rorschach study of a group of scientists and technicians. *J. consult. Psychol.*, 1946, 10, 317-327.
- Roe, Anne. Analysis of Group Rorschachs of biologists. *J. proj. Tech.*, 1950, 14, 384-398.
- Roe, Anne. Analysis of Group Rorschachs of psychologists and anthropologists. *J. proj. Tech.*, 1952, 16, 212-224.
- Roe, Anne. *The psychology of occupations*. New York: John Wiley and Sons, 1956.
- Roe, Anne. Early determinants of vocational choice. *J. counsel. Psychol.*, 1957, 4, 212-217.

Teevan, R. C. Personality correlates of undergraduate field of specialization. *J. consult. Psychol.*, 1954, 18, 212-214.

Interest Inventory scores and other attitudes and personality factors. *J. appl. Psychol.*, 1945, 29, 58-67.

Tyler, Leona E. Relationships between Strong

Received December 12, 1961

Some Normative Data on Rorschach Developmental Level "Card Pull" in a Psychiatric Population

MARVIN R. GOLDFRIED¹
University of Rochester

In recent years, several investigators (DeVos, 1952; Elizur, 1949; Fisher & Cleveland, 1958; Friedman, 1953; Holt, 1956) have been concerned with developing special Rorschach scoring systems. One such system, developed by Friedman (1953), was devised to assess developmental level of perceptual organization. This developmental scoring of Rorschach responses is based in large part on Werner's theory of cognitive and perceptual development (Werner, 1948), which states that development proceeds from a lack of differentiation to a more articulated and integrated mode of functioning. Consequently, the Rorschach scoring categories formulated by Friedman parallel the various stages in this developmental process.

Friedman's scoring system differs from the conventional Rorschach scoring in that only the structural and organizational aspects of the percept are used. Friedman has used location scores only, classifying the responses according to the level of diffuseness, articulation, integration, etc. Thus, genetically low scores include *amorphous* (where form is not used as a determinant) whole (W) and detail (D) responses, *vague* responses (where form "is of such an unspecified nature that almost any perceptual form is adequate"), and *minus* responses (which are poor fits to the stimulus properties of the blot). Genetically high scoring categories applied to W and D responses are *plus plus*, *plus*, and *mediocre*. A *plus plus* score is one in which a single area of the blot "is perceptually articulated and then re-

integrated into a well differentiated" unified response of adequate form level. In the case of W responses, a *plus plus* score (W++) can be given only to what Friedman has called "unbroken blots" (Cards I, IV, V, VI, and IX). A *plus* score is given to those responses of adequate form quality where two separate areas have been combined into a single percept "Broken blot" (Cards II, III, VII, VIII, and X) are the only cards to which a W+ score can be applied. A *mediocre* scoring is applied to responses of adequate form level that are given to a single area; hence, a Wm score can occur only to unbroken blots.

The inter-scorer reliability of this system is high, ranging from 89.7 per cent to 95.5 per cent agreement between judges (Friedman, 1955). Studies pertaining to the validity of the scoring system also have been encouraging. In Friedman's study, it was found that children have significantly more genetically low than high responses; with normal adults, the converse was true. A study by Hemmendinger (1953) found that genetically low scores decreased and genetically high scores increased between the ages of 3 and 10. Siegel (1953) tested and confirmed the hypothesis that catatonic and hebephrenic schizophrenics are more regressed than paranoid schizophrenics, finding the latter group to give percepts of a higher developmental level. Studies also have found that "reactive" schizophrenics have a higher developmental level than "process" schizophrenics (Becker, 1956; Fine and Zimet, 1959). In general, the results of these studies indicate that the scoring system represents a valid approach to the assess-

¹The author is indebted to the staffs of the V.A. Neuropsychiatric Hospital in Palo Alto, California and the V.A. G.M. & S. Hospital in Buffalo, N.Y. for making the Rorschach protocols available.

TABLE I—Frequency and Per Cent of Occurrence of Each Scoring Category for Each Rorschach card (N=110)

Level	Scoring Category	I	II	III	IV	V	Card	VI	VII	VIII	IX	X	Occurrence Of Each
		Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	Freq. %	
1	Amorphous Whole (Wa)	0 0.00	0 0.00	0 0.00	1 0.53	0 0.00	0 0.00	0 0.00	5 2.86	3 1.49	4 2.07	3 0.86	0.73
	Minus Whole (W-)	15 5.86	13 5.94	2 0.87	11 5.82	15 8.02	3 1.59	17 9.71	17 9.71	9 4.46	8 4.15	5 1.43	4.47
	Contributory Response (DW)	0 0.00	3 1.37	1 0.43	1 0.53	1 0.53	0 0.00	0 0.00	0 0.00	1 0.50	0 0.00	0 0.00	0.36
	Contaminated Response (ConR)	0 0.00	1 0.46	1 0.43	0 0.00	1 0.53	0 0.00	2 1.14	0 0.00	0 0.00	0 0.00	1 0.29	0.27
	Fabulized Combination (FabC)	1 0.39	1 0.46	1 0.43	2 1.06	5 2.67	5 2.65	0 0.00	0 0.00	1 0.50	0 0.00	0 0.00	0.75
	Rejection ^a	0 0.00	2 0.91	1 0.43	6 3.17	4 2.14	10 5.29	7 4.00	1 0.50	1 0.50	10 5.18	5 1.43	2.10
2	Amorphous Detail (Da)	0 0.00	0 0.00	1 0.43	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	2 1.00	6 3.11	2 0.57	0.50
	Contributory Detail (DdD)	0 0.00	0 0.00	1 0.43	0 0.00	1 0.53	1 0.53	1 0.57	0 0.00	0 0.00	2 1.04	1 0.29	0.32
	Minus Detail (D-)	12 4.69	30 13.70	29 12.55	8 4.23	6 3.21	13 6.88	12 6.86	13 6.44	22 11.40	52 14.90	25 7.16	8.99
	Vague Detail (Dv)	9 3.52	8 3.65	10 4.33	3 1.59	2 1.07	10 5.29	4 2.29	16 7.92	7 3.63	25 7.16	7 1.96	4.29
3	Minus Unusual Detail (Dd-)	4 1.56	6 2.74	2 0.87	11 5.82	7 3.74	5 2.65	1 0.57	2 1.00	8 4.15	0 0.00	0 0.00	2.10
	Vague Whole (Wv)	13 5.08	6 2.74	3 1.30	11 5.82	8 4.28	7 3.70	28 16.00	17 8.42	25 12.95	14 4.01	14 4.01	6.03
	Oligophrenic Detail (Adx-Hdx)	0 0.00	0 0.00	3 1.30	0 0.00	2 1.07	1 0.53	0 0.00	0 0.00	0 0.00	1 0.52	0 0.00	0.32
	Plus Unusual Detail (Dd+)	33 12.89	17 7.76	4 1.73	7 3.70	10 5.35	12 6.35	6 3.43	6 2.97	23 11.92	5 1.43	5 1.43	5.62
4	Mediocre Detail (Dm)	55 21.48	58 26.48	85 36.80	43 22.75	26 13.90	70 37.04	48 27.43	85 42.08	49 25.39	171 49.00	31.51	
	Mediocre Whole (Wm)	97 37.89 ^b	69 36.51	91 48.66	45 23.81	2 ^c 1.14	4 2.07	14.06	
5	Plus Detail (D+)	4 1.56	41 18.72	63 27.27	0 0.00	0 0.00	2 1.06	14 8.00	26 12.87	11 5.70	48 13.75	9.54	
	Plus Whole (W+)	33 15.07	24 10.39	28 16.00	20 9.90	15 4.30	5.48	
6	Plus-Plus Detail (D++)	4 1.56	0 0.00	0 0.00	2 1.06	2 1.07	0 0.00	0 0.00	0 0.00	0 0.00	2 1.04	2 0.57	0.55
	Plus-Plus Whole (W++)	9 3.52	14 7.41	6 3.21	4 2.12	11 5.70	2.00	
Mean no. Responses		2.3	2.0	2.1	1.7	1.7	1.6	1.5	1.5	1.8	1.7	3.1	
Mean Developmental Level		3.55	3.56	3.88	3.51	3.39	3.34	3.31	3.36	3.13	3.13	3.56	

^a—The classification of rejections at level 1 has been suggested by Wilemsky (1959b)^b—This indicates that the particular score is not applicable to the card, either because the blot is "broken" or "unbroken"^c—While Wm cannot usually be applied to a "broken blot", it can occur on card VII if the response implies a definite "U" shape, such as, "harbor," "bowl," etc. Friedman, 1953

TABLE II Per Cent of Mature and Immature Scoring Categories for all Whole (W) and Usual Detail (D) Locations

	Mature			total % mature	Immature			Confab. Response (DW or DdD)	Total % Immature
	++	+	m		-	v	a		
W	6.1	16.5	42.4	65.0	13.5	18.2	2.2	1.1	35.0
D	1.0	17.1	56.6	74.7	16.1	7.7	0.9	0.6	25.3

Note: The percentages are based on the proportion of the scoring category to location score in general (e.g. $W++/\text{Total } W$, $W+/\text{Total } W$, etc.).

ment of developmental level of perceptual organization.²

The purpose of the present investigation is to present some normative data on Rorschach developmental level in a psychiatric population. An attempt was made to determine the developmental level "card pull" for each of the Rorschach cards; that is, the level of perceptual organization usually elicited by each card.

METHOD

Rorschach protocols were obtained from 110 male neuropsychiatric patients in two V.A. hospitals. Fifty patients were at an N.P. hospital, while 60 were hospitalized on an N.P. ward in a G. M. & S. hospital. The mean age of all 110 patients was 34.3, with an age range of 18-57 years. I.Q. scores were available on only 89 patients in the sample; of these tested (either by WAIS or Wechsler-Bellevue), the mean I.Q. was 104.9, with a range of 79-129.

The Rorschach protocols were all rescored according to Friedman's (1953) developmental scoring system. Using Becker's (1956) rankings of the scoring categories (see Table I), an overall Development Level (*DL*) was obtained for each of the 10 Rorschach cards. Further, the per cent of occurrence of each scoring category on each card, and in the sample in general was determined. Comparison was also made between two methods

of computing *DL* (Becker, 1956, Wilensky, 1959a).

RESULTS AND DISCUSSION

Analysis of data

The frequency and the per cent of occurrence of each scoring category for each card is presented in Table I. The scoring categories are arranged in the table according to Becker's (1956) ratings of their "maturity." The omission of data at levels 4, 5, and 6 is due to the fact that *Wm* and *W++* can occur only on "unbroken blots", and *W+* only on "broken blots."

The mean *DL* for each card was computed by giving each response a weight of one through six as indicated in Table I, adding all these scores, and dividing by the number of responses (Becker, 1956). As can be seen in Table I, the blot obtaining the highest *DL* was Card III. The scoring categories of *D+* and *W+*, which helped to contribute to this relatively high *DL*, usually were achieved by responses to this card which depicted two people doing something. The blot with the lowest *DL* was Card IX, which is consistent with the usual clinical observation that this card is the most difficult to organize perceptually of all the cards.

Table II summarizes some of the data in a way that may be more useful to routine clinical usage. This table presents the per cent of each *W* and *D* scoring category to all *W* and *D* scores respectively. The scoring categories are grouped according to Friedman's (1953) distinction between "mature" and "immature" scores. Thus, the categories which most fre-

²The possibility that developmental level of perceptual organization may reflect a different level of functioning than tests such as the MMPI is discussed elsewhere (Goldfried, 1962).

quently contribute to the "maturity" of the Rorschach protocols in this sample are *Wm* and *Dm*. The "immature" aspects of the records are composed mainly of *Wv*, which is believed to be an indication of regression (Rosenblatt & Solomon, 1954; Seigel, 1953), as well as *W-* and *D-*, which typically are interpreted as reflecting an impairment of reality testing.

Comparison of two methods of computing DL

Wilensky (1959a) has proposed an alternate method of obtaining the overall *DL* score from the Rorschach protocol. According to Becker's approach, the weights of one through six are summed for all cards, which is then divided by the total number of responses to obtain the *DL*. Wilensky, while he uses the same weights as Becker does, totals the score for each card separately and divides by the number of responses for that card. In other words, he gets a *DL* for each card, and averages all 10 *DL* scores to obtain an overall *DL*. Wilensky maintains that calculating *DL* by his method has the advantage of minimizing the effect that certain cards have on the overall *DL*. Thus, performance on Card X may spuriously increase *DL* (as computed by Becker's method), as this card tends to elicit many *Dm* responses (see Table I). Wilensky further argues that his method of computing *DL* also serves to increase the variability of scores, which is particularly important when one is working with a fairly homogeneous group.

In order to compare Becker's and Wilensky's methods of computation, the *DL*'s for all 110 protocols in the present study were calculated by both methods. The mean *DL*'s and standard deviations of the two methods are presented in Table III. Wilensky (1959a) reports a correlation of .95 between the two methods. The correlation between the Becker and Wilensky *DL*'s obtained in the present study is .89 ($p < .001$). To test whether the Wilensky method produces a more variable distribution, the variances

TABLE III—Means and S.D.'s of *DL* as Calculated by Becker's and Wilensky's Methods

	Becker's <i>DL</i>	Wilensky's <i>DL</i>
Mean	3.50	3.75
S.D.	0.412	0.485

were compared by means of an *F* test. The results of this comparison confirm the fact that Wilensky's *DL* does produce greater variability ($F=1.38$, $df=109$ and 109 , $p < .05$).

The relationship between DL and I.Q.

Viewing Rorschach *DL* as reflecting the individual's more general ability to function cognitively, one might expect to find some relationship between *DL* and I.Q. Ainsworth & Klopfer have indicated that "vague, global perception reflects a relatively low level of capacity, and . . . the more refined and differentiated the perception the higher the level of intelligence" (Ainsworth & Klopfer, 1954, p. 353). To test this hypothesis, Rorschach *DL* was correlated with I.Q. scores for those 89 cases where I.Q. measures were available. The correlation between Becker *DL* and I.Q. was .39 ($p < .001$); for Wilensky *DL*, the correlation achieved was .44 ($p < .001$). As the Wilensky method produces a more variable distribution of *DL*, one might have expected to obtain a higher correlation in comparison with Becker's *DL*. The actual statistical test of the differences between these two *r*'s (Edwards, 1960), however, failed to reach significance ($t=1.09$). While a correlation has been found to exist between Rorschach *DL* and I.Q. in a psychiatric population, further empirical evidence is needed to determine the relationship between these two variables for normal Ss.

Research uses

In addition to the clinical use that the normative data presented in Tables I and II can be put to, the *DL* "card pull" may be useful for research on developmental level of functioning as well. For example, knowledge of the characteristic *DL* for each card enables

investigators to use half the cards in each of two experimental conditions and still maintain comparable DL "card pull." This allows for the use of the same S in each condition and reduces the variability due to individual differences. Further, by constructing two sets of five Rorschach cards each which are comparable as to DL "card pull", one eliminates the effects that memory of original responses may play in a test-retest design. The use of the two sets therefore can facilitate what Werner (1948) has called the "genetic experiment;" that is, the experimental alteration of an individual's developmental level of functioning. For example, based on the hypothesis that ego regresses to earlier stages of development under such conditions of exhaustion, intoxication, sleepiness, etc. (Fenichel, 1945), one might study the changes that these variables produce in Rorschach developmental level.

REFERENCES

- Ainsworth, Mary D., & Klopfer, B. Evaluation of intellectual level, control, creative potential, and introverted-extraverted relationship. In Klopfer, B. *et al. Developments in the Rorschach technique*. Vol. I. Yonkers-on-Hudson, N. Y.: World Book, 1954.
- Becker, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, 53, 229-236.
- DeVos, G. A quantitative approach to affective symbolism in Rorschach responses. *J. proj. Tech.*, 1952, 16, 133-150.
- Edwards, A. L. *Experimental design in psychological research* (Rev. ed.), New York: Rinehart, 1960.
- Elizur, A. Content analysis of the Rorschach with regard to anxiety and hostility. *J. proj. Tech.*, 1949, 13, 247-284.
- Fenichel, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- Fine, H. J., & Zimet, C. N. Process-reactive schizophrenia and genetic levels of perception. *J. abnorm. soc. Psychol.*, 1959, 59, 83-86.
- Fisher, S., & Cleveland, S. E. *Body image and personality*. Princeton, N. J.: Van Nostrand, 1958.
- Friedman, H. Perceptual regression in schizophrenia: an hypothesis suggested by use of the Rorschach Test. *J. proj. Tech.*, 1953, 17, 171-185.
- Goldfried, M. R. Rorschach developmental level and the MMPI as measures of severity of psychological disturbance. *J. proj. Tech.*, 1962, in press.
- Hemmendinger, L. Perceptual organization and development as reflected in the structure of the Rorschach test responses. *J. proj. Tech.*, 1953, 17, 162-170.
- Holt, R. R. Gauging primary and secondary processes in Rorschach responses. *J. proj. Tech.*, 1956, 20, 14-25.
- Rosenblatt, B., & Solomon, P. Structural and genetic aspects of Rorschach responses in mental deficiency. *J. proj. Tech.*, 1954, 18, 496-506.
- Siegel, E. L. Genetic parallels of perceptual structuralization in paranoid schizophrenia: an analysis by means of the Rorschach technique. *J. proj. Tech.*, 1953, 17, 151-161.
- Werner, H. *Comparative psychology of mental development*. (rev. ed.), Chicago: Follet, 1948.
- Wilensky, H. Rorschach developmental level and social participation of chronic schizophrenics. *J. proj. Tech.*, 1959, 23, 87-92 (a).
- Wilensky, H. Developmental scoring of Rorschachs of schizophrenics. Paper presented at East. Psychol. Assoc., Atlantic City, April, 1959 (b).

Received March 16, 1962

Symbol Consensus and Univocality in Schizophrenia¹

ALFRED E. GOLDMAN
National Analysts, Inc.²

That symbolic functioning is seriously affected in schizophrenia has been long acknowledged by clinical observers (Arieti, 1948; Kasanin, 1944; Kubie, 1953). However, it has remained to document this impairment and to specify its nature and genesis.

The earlier notion that linguistic functions were the least likely to be affected in schizophrenia was based on the observation that the level of performance in vocabulary subtests of intelligence tests was maintained relative to performance in other spheres (Babcock, 1930). Subsequent studies have indicated convincingly that schizophrenics respond to the vocabulary items at a qualitatively different level of functioning from normals (Richman, 1957; Yacorzinski, 1941). These studies highlight the distinction between achievement and process and make clear that the vocabulary performance "holds" in pathology by virtue of a test artifact, and reflects scoring criteria which focus on achievement rather than the underlying process. The goal of understanding psychic processes demands a different method from one which categorizes behavior in terms of success and failure (Goldstein & Scheerer, 1941; Werner, 1937).

This study proceeds from the recognition of a dysfunction in the schizophrenic's symbolizing activity (Cameron, 1938, 1939; Arieti, 1948). The aspect of symbolic functioning which is most apparent at the interpersonal level is the degree to which members of the same group agree in the meaning they give words. This is what is

meant by *consensus*. The first factor to be considered here is the degree to which schizophrenics and normals agree, among themselves, as to the meaning given words (*consensus*). Consonant with a comparative developmental point of view (Werner, 1940, 1946), which is the theoretical context within which this study was conceived³, hypothesis one is that agreement in meaning among schizophrenics will be lower than among normals.

A necessary prior condition for agreement in symbol-referent relationships between people is the stability of this relationship within each of the individuals that comprise a group. This stability between a word and its meaning is called *univocality*. A certain level of consistency in the meaning ascribed to a word must be achieved in order for that word to be public, or socially useful. Thus, a second hypothesis is that uniformity in the meaning of words within a given individual (*univocality*) will be lower for schizophrenics than for normals.

PROCEDURE

Exploration of the meaning of words raises special problems with respect to method. Inherent in verbal methods, in which subjects "tell" the examiner what they understand the meaning of a word to be, is the problem of methodological circularity. The very vehicle under study is used as an exploratory device. To avoid this problem a method was sought in which meaning was mediated by other than verbal means. Of a variety of techniques available (Osgood, 1953), the linear-schematization technique (Immersgluck & Cohen in Krout, 1950; Hall & Oldfield, 1950; Scheerer &

¹Based in part on data from a doctoral dissertation completed at Clark University. The author is indebted to Drs. Heinz Werner, Bernard Kaplan, and Seymour Wapner for their aid in this study.

²1015 Chestnut Street, Philadelphia, 7, Pa.

³This developmental theoretical context is described elsewhere (Goldman, 1960, 1962).

Lyons, 1957; Pollack, 1953; Goldman, 1955) was thought to be the best suited. In the linear-schematization technique the subject is required to express the meaning of a word by drawing any line or shape. This technique affords the following three advantages:

a. It permits the expression of the subjectively experienced meaning of a word which is not ordinarily reflected in codified language. Lines, in a manner similar to that of bodily gestures, convey a level of personal experience that codified language conveys only when very skillfully employed as the expressive medium. The technique, for this purpose, is better than "provoked introspection" in that it does not require of the subject any unusual verbal skill. The lines allow the subject to transcend both his own expressive inadequacy and the limitations of the linguistic code.

b. It permits the articulation of meaning in an objective form which is therefore available to intersubjective observation. Thus externalized, it can be classified and quantified.

c. It serves to stabilize a transitory experience. Our hypothesis, dealing with lability of meaning, requires that the technique employed be able to "freeze" the particular manner in which a subject experiences a meaning at a given moment in time, so that any change in that meaning is available to inspection.⁴

While the linear-schematization technique may provide an index of meaning, the quantitative analysis of the structural criteria of the lines is cumbersome. The problem of classifying into fixed categories and quantifying the line responses of normals (Scheerer & Lyons, 1950) is substantially compounded when the lines are produced by schizophrenics.

To circumvent this problem a method was sought which would permit strict quantitative analysis of the formal criteria of differentiation of the

lines. The method of response-equivalence proved to be well suited to this purpose. The task consists of having the subject match one of several lines with one word. This task situation is a modification of a prior response-equivalence technique developed by the author (Goldman, 1955).

The choice of emotionally-toned terms as stimulus words was determined by several factors.

First, affective words are easier to dichotomize into polar opposites than neutral words. Secondly, if an impairment in the social use of words exists it is more likely to be manifest with emotionally charged words than with affectively neutral material.

The response-equivalence technique used in this experiment differs from that used in the prior experiment (Goldman, 1955) in one important respect, and was designed to achieve greater precision in relating the structural characteristics of lines with mood-terms. In the first experiment the stimulus cards presented two words — polar opposites, such as love and hate — and one line. The line was deliberately overdetermined in its structural characteristics so that lines that were hypothesized to be "negative" in affective value were drawn very dark, angular, and in a downward direction. Similarly, when the same two words appeared later in the series of stimulus cards with a single line having antithetical structural characteristics, this line too, combined three supposedly positive characteristics: lightness, circularity, and drawn in an upward direction. Thus, in pairing words with lines the subjects may have been responding to any one or more of the line characteristics included.

Since the results of the first experiment indicated that normals produce significantly more negative word-negative line and positive word-positive line matchings than schizophrenics, a second experiment was suggested in which the structural characteristics of the lines would be independently

⁴See B. Kaplan's discussion of the advantages of this technique in Werner (1955).

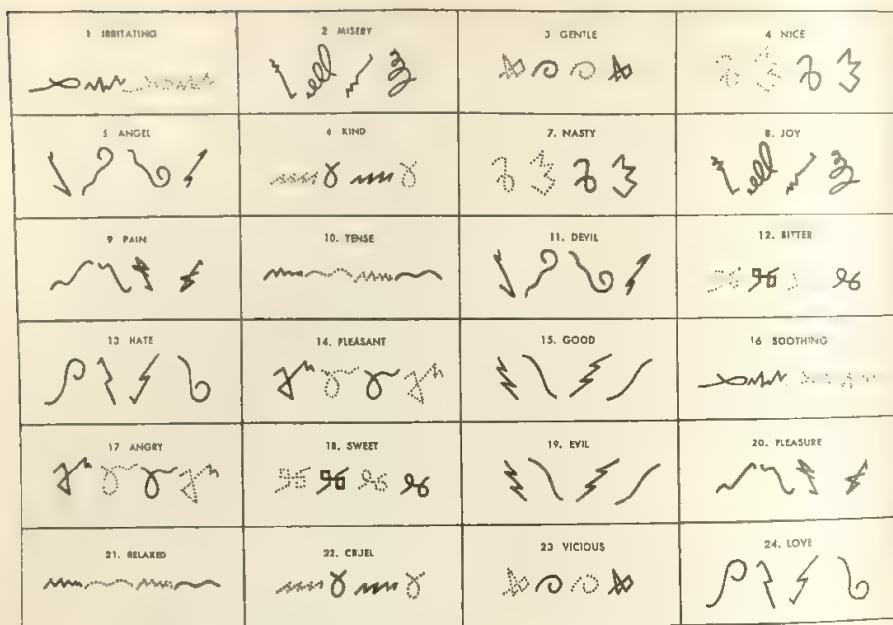


FIGURE 1—Stimulus Material (Dots represent light lines)

varied.

Before discussing the advantages of this modification it is necessary to present the procedure itself. In the present experiment subjects were shown twenty-four stimulus cards, each with one word and four lines. Each four-line set appears twice in the series with opposite members of a mood-word pair; for example, a four-line set that first appears with an affectively positive mood word (joy) appears later in the series with its opposite (misery). The following three characteristics of the stimulus lines are systematically varied and presented in random order (see Figure 1 for order of presentation).

1. Angularity-circularity: On each card two of the lines are angular and two circular.

2. Direction: On ten cards, two lines are drawn in an upward direction (one angular, one circular) and two lines are drawn downward (one angular, one circular).

3. Darkness:⁵ On fourteen cards,

⁵ More precisely—darkness and thickness. The

two lines are dark (one angular, one circular) and two light (one angular, one circular).

The following instructions were given: "Here is a word and some lines. Choose the line that you think goes best with this word — the line that best represents the word." Following response to card 2, 6, 8, and 22,⁶ subjects were asked: "What is there about this line (the line which was selected as best representing the one word) that suggests the word——?" Line choice and inquiry were recorded on protocol sheets.

The advantages of this modification of the response-equivalence technique were twofold. First, it permits independent exploration of the three line characteristics, angularity, darkness and direction, in terms of discriminating the performance of schizophrenics and normals. Secondly, and more im-

portant, dark lines in this experiment were drawn with a very black, soft lead pencil which produced a line that was thicker than that of the light line. Thus, the darkness of the "dark" lines was accentuated considerably.

⁶ Arbitrarily selected.

portant, this modification allows for analysis in terms of degrees of univocality. Since each four-line set appears twice in the twenty-four card series, once with a positive word, once with a negative word, three selection alternatives are available to the subject.

Alternative 1. Choice of lines sharing no common characteristics for each of the opposite mood-words.

Alternative 2. Choice of lines sharing one common characteristic for each of the opposite mood-words.

Alternative 3. Choice of the identical line, sharing two common characteristics, for each of the opposite mood-words.

In the case of the first alternative, in which opposite words share no common line characteristics, it appears that these words are fully dichotomized with respect to the expression of meaning. In our sense these words may be said to be semantically *univocal*. It follows then, that opposite mood-terms which share one common line characteristic are less differentiated expressively, and thereby relatively more unstable (*equivocal*) semantically than opposite words which are fully separated in their symbolic representation. Finally, opposite words which are expressed by the identical line (i.e., two common line characteristics) may be said to be the most *equivocal* with respect to the expression of meaning. In accordance with the second hypothesis we expect that schizophrenics will link opposite words, through selection of both one and two common characteristics, more frequently than normals. Insofar as the selection of identical lines to represent opposite terms reflects greater equivocality, we may also expect this criterion of equivocality to distinguish the two groups more sensitively than linkage of antonymic terms by one common characteristic.

Subjects

Subjects were 15 normal and 15 schizophrenic women. The normal women were all members of some social or charitable organization or

woman's auxiliary, and none had ever been hospitalized or treated for emotional illness. The schizophrenic sample was drawn from the female patient population of Boston State Hospital and had the following characteristics:

(a) Each manifested symptoms which were unequivocally psychotic, and had been diagnosed schizophrenic by at least two psychiatrists; (b) Each could fulfill at least the minimal requirements of the experiment. Patients whose attention drifted or who could not cooperate for other reasons were not used as subjects.

Subjects were individually matched in pairs for years of education. A t-test indicated no statistically significant differences exist between the normal and schizophrenic samples for educational level and age.

RESULTS AND DISCUSSION

Degree of Consensus (agreement among subjects)

The relative consensus with which the two groups expressed meaning was evaluated by the degree to which each group selected the modal choice of the group of which he was a member. Thus, in order to evaluate hypothesis 1 these modal choices had to be determined. These data are presented in Table I. The reliability of these data

TABLE I—Number of Subjects Choosing Majority of Line-Characteristics for Mood-Terms Corresponding to Assumed Conventional Relationships

	Normals (N=15)		Schizophrenics (N=15)	
	No. of Sub- jects	p	No. of Sub- jects	p
Circular line chosen				
for "positive" word	15	<.001	11	<.05
Angular line chosen				
for "negative" word	15	<.001	12	<.05
Light line chosen				
for "positive" word	13	<.01	9	>.05
Dark line chosen				
for "negative" word	15	<.001	12	<.05
Upward line chosen				
for "positive" word	15	<.001	11	<.05
Downward line chosen				
for "negative" word	13	<.01	8	>.05

were evaluated by means of the statistical sign test (Dixon, 1946; Mosteller, 1954).

Reference to Table I indicates that normals, in the majority of trials, match circular lines, light lines, and upward lines to positive mood-terms, while angular lines, dark lines and downward lines are matched with negative mood-terms. These relationships are found to be reliable at or below the .01 level of significance. The modal line characteristic and mood-term matchings are the same for the schizophrenics as for the normals. That is, in the majority of trials, most schizophrenics select circular, light and upward lines to represent positive words, and most schizophrenics select angular, dark and downward lines to represent negative words. For the schizophrenics, however, the relationship between light line and positive word, and the relationship between downward line and negative word fails to meet the .05 criterion of statistical significance.

In summary, the data provided by both groups lend credence to the assumption, derived from the literature, relating line characteristics to affective valence. In accordance with our operational definition of consensus, we may state that for both the normal and schizophrenic groups, the intuitive meaning of the affectively positive words are *consensually* expressed in terms of circularity, lightness, and upwardness of line, while the intuitive meaning of affectively negative words are consensually expressed in angularity, darkness, and downwardness of line.

Test of the expectation that agreement in the expressions of meaning is lower among schizophrenics than among normals inquires into the relative degree to which members of each group conform in their matchings to the modal choices of their own group. The significance between the distribution of consensual choices for the two groups is evaluated by means of the Kolmogorov-Smirnov c' statistic, which tests the null hypothesis that the cumulative distribution of values from one sample is equal to, or lower than (one tail test), the cumulative distribution of values from a second sample (Goodman, 1954; Massey, 1951; Massey, 1951a).

Considering the structural criteria independently (Table II) for each criterion, normals exceeded schizophrenics in their matching of the consensual line-word relationships. For five of the six relationships tested, the differences between the two groups were at or below the .05 level of significance. Although the relationship between downward lines and negative words was not significantly different in terms of statistical convention, the same tendency was found.

Degree of Univocality (stability of choice within subjects)

In accordance with hypothesis two, greater equivocality in the expression of meaning was expected among schizophrenics than normals. The differences between the normal and schizophrenic groups with respect to the distribution of equivocal responses was evaluated by means of the Kolmogorov-Smirnov c' statistic.

TABLE II—Degree of Consensus: Mean Number of Consensual Choices

	Normal	Schiz.	c'	p
Circularity-Angularity (12 trials)				
Positive Word—Circular Line	11.60	8.53	$8\frac{1}{15}$	$<.05$
Negative Word—Angular Line	11.53	8.40	$13\frac{1}{15}$	$<.001$
Lightness-Darkness (7 trials)				
Positive Word—Light Line	6.20	4.20	$7\frac{1}{15}$.05
Negative Word—Dark Line	6.80	5.20	$7\frac{1}{15}$.05
Upward-Downward (5 trials)				
Positive Word—Upward Line	4.40	3.00	$8\frac{1}{15}$	$<.05$
Negative Word—Downward Line	3.73	2.87	$5\frac{1}{15}$	$>.05$

TABLE III Degree of Univocality: Mean Number of Choices of Common Line Characteristics for Opposite Words

	Normal	Schiz.	χ^2	P
One Common Characteristic	2.40	1.27	1.73	.09
Two Common Characteristics (Identity)	.07	1.73	1.73	.001

In this analysis, focus was on the three alternatives permitted the subject: choice of lines sharing no common characteristics, one common characteristic, or the identical line (two common characteristics) for each of the opposite mood-word pairs. The expectation of greater equivocality for the schizophrenics appear to be confirmed by the results presented in Table III. The schizophrenics selected, for opposite mood-terms, significantly more lines having one common characteristic than did the normals. The differences between the two groups with respect to choice of identical lines to represent opposite mood-terms is particularly striking. Here we find thirteen schizophrenics contributing a total of 25 identical line choices for opposite words (mean=1.73), while only one of the fifteen normal subjects yielded one such response (mean=.07).

Insofar as an individual can express the meaning conveyed by opposite mood-terms with the identical expressive line, then it would seem that these words are not differentiated as symbols. In his use of the word, then, the individual can call upon one or another meaning, dependent upon the needs of the moment.

Similar symbol equivocality has been found in the language of primitive societies (Werner, 1940), in the drawing behavior of children (Eng, 1931), in dreams (Freud, 1950), and in archaic languages (Freud, 1950), and thus suggests its developmental primitivity.

SUMMARY

This study investigates two characteristics of symbolic functioning in schizophrenics—agreement in meaning between people (consensus) and stability in meaning within any one

person (univocality). The method requires subjects to relate lines with mood-words. The results indicate that schizophrenics, as compared to normals, are significantly less consensual and less stable in their expression of meaning.

REFERENCES

- Arieti, S. Special logic of schizophrenia and other types of autistic thought. *Psychiatry*, 1948, 11, 325-338.
- Babcock, H. An experiment in the measurement of mental deterioration. *Arch. of Psychol.*, 1930, 18, No. 117.
- Cameron, N. Reasoning, regression and communication in schizophrenics. *Psychol. Monogr.*, 1938, 50, 221.
- Cameron, N. Schizophrenic thinking in a problem solving situation. *J. Ment. Sci.*, 1939, 1-24.
- Dixon, W. J., & Mood, A. M. The statistical sign test. *J. Amer. Statist. Assoc.*, 1946, 41, 557-566.
- Eng, H. *The psychology of children's drawings*. London: Kegan Paul, 1931.
- Freud, S. The antithetical sense of primal words. In E. Jones (ed.), *Collected Papers*, Vol. IV, London: Hogarth Press, 1958.
- Goldman, A. E. Symbolic representation in schizophrenia. *J. Pers.*, 1960, 28, 293-316.
- Goldman, A. E. A Comparative-Developmental Approach to Schizophrenia. *Psychological Bulletin*, 1962, 59, 57-69.
- Goldstein, K. & Scheerer, M. Abstract and concrete behavior. *Psychological Monogr.*, 1941, 53 (Whole No. 239).
- Goodman, L. A. Kolmogorov-Smirnov tests of psychological research. *Psychological Bulletin*, 1954, 53, 2, 160-168.
- Hall, K. R. L., & Oldfield, R. C. An experimental study of the fitness of signs to words. *Quart. J. exp. Psychol.*, 1950, 2, 60-70.
- Immersgluck & Cohen. In Krout, J. Symbol elaboration test: the reliability and validity of a new projective technique. *Psychological Monogr.*, 1950, 64, 310.
- Kasanin, J. S. *The disturbance of conceptual thinking in schizophrenia*. Berkeley and Los Angeles: University of California Press, 1944.
- Kubie, L. S. The distortion of symbolic process in neurosis and psychosis. *J. Amer. Psychoanal. Assoc.*, 1953, 1, 59-86.
- Massey, F. J. Jr. The distribution of the maximum deviation between two sample cumu-

- lative step-functions. *Ann. Math. Statist.*, 1951, 22, 125-128.
- Massey, F. J. Jr. The Kolmogorov-Smirnov test for goodness of fit. *J. Amer. Statist. Assoc.*
- Mosteller, F., & Bush, R. Selected quantitative techniques. In Lindzey, G. (Ed.), *Handbook of Social Psychology*. Cambridge, Massachusetts: Addison-Wesley, 1954.
- Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford University Press, 1953.
- Pollack, R. H. A genetic study of intuitive word meanings. Unpublished Ph.D. dissertation, Clark University, Worcester, Massachusetts, 1953.
- Richman, J. The effect of the emotional tone of words upon the vocabulary responses of schizophrenics. *J. gen. Psychol.*, 1957, 56, 95-119.
- Schacter, M., & Lyons, J. Line drawings and matching responses to words. *J. Pers.*, 1957, 25, 251-273.
- Werner, H. Process and achievement — a problem of educational and developmental psychology. *Harvard Education Review*, 1937, 7.
- Werner, H. *Comparative psychology of mental development*. New York: Harper, 1940.
- Werner, H. Genetic experimental psychology. *Encyclopedia of Psychology*. P. 1. Thurman (Ed.). Philosophical Library, 1946.
- Werner, H. (Ed.). *On expressive language*. Worcester, Massachusetts: Clark University Press, 1955.
- Yacorzynski, G. K. An evaluation of the postulates underlying the Babcock deterioration test. *Psychol. Rev.*, 1951, 48, 261-267.

Received October 30, 1961

The Role of Diagnostic Evaluation in Clinical Psychology

WALTER G. KLOPPER
University of Portland

It has become fashionable during recent years among some clinical psychologists to derogate the diagnostic evaluation as a professional activity. The purpose of the present article is to examine the various reasons given for this lowering of diagnostic evaluation on the hierarchy of professional activities and to see whether they will bear close scrutiny. Thus, cited below are the six major arguments given on behalf of the pessimistic position, each argument being followed by a critical evaluation.

1. "Diagnostic testing is a relatively simple activity which requires less skill and training than other responsibilities of the clinical psychologist, such as psychotherapy and research."

Recent reviews by Masling (1959) and Guertin *et al* (1962), have demonstrated that psychological tests are not the simple laboratory procedures they are sometimes assumed to be. Not only are projective instruments subject to transient and situational influence, but even such seemingly objective instruments as intelligence tests can be manipulated while remaining within the standard administrative regulations. In order to interpret psychological tests properly it is necessary not only to have a great deal of clinical skill and experience, but also to be keenly aware of the limitation of predictive efficiency imposed upon the clinician by the normative data upon which predictive hypotheses are based, the level of awareness being tapped, the previous experience of the patient in taking tests, his current motivation and set, and of course, the biases, predispositions, and interpersonal relationships of the clinical psychologist doing the test interpreting. Viewed in this light the process becomes considerably more complex. It is the firm conviction of this author that nowhere

does the clinical psychologist utilize more of his resources than in the area of diagnostic evaluation. Psychotherapy, for all its complex theoretical bases, is still practiced mainly as an art at the functional level. Research can be ingenious and creative, but is often carried out at the hack level. However, to do an effective job of diagnostic evaluating requires the use of all the clinical psychologist's scientific and clinical training, plus the best intuition he is capable of bringing forth.

2. "Diagnostic evaluation, although of value, is not as worthy of taking up the psychologist's time as other activities of greater direct service to society, such as psychotherapy."

Of course there are many different ways of doing psychotherapy. The client-centered therapist has maintained stoutly that having available facts about the patient, whether these be anamnestic data or psychological test data, is of little value since the therapist does not plan the treatment but merely provides an atmosphere within which the patient can mature and grow on his own. It is outside of the scope of this paper to examine this particular assumption in detail, except to point out that only some patients can progress under these circumstances, and that it may be important to determine in advance which patients may or may not benefit from such a procedure. In order to answer this question, a diagnostic evaluation is a prime requisite.

Another method of psychotherapy, the orthodox psychoanalytic one, operates mainly at the intra-personal or intra-psychic level. Thus the patient is in effect asked to shake hands with his own unconscious and to gain conscious control over his various motives so that he may not be overwhelmed in

the future and his ego weakened by stresses from within. This seems to necessitate at least some diagnostic evaluation, even though this evaluation can be limited to intrapersonal types of evaluations such as are carried on by means of projective techniques.

The third and more currently common approach to psychotherapy embraces a realistic evaluation on the part of the therapist of the entire phenomenological world of the patient. This includes his inner stresses, his adaptation to reality, his characteristic interpersonal methods, and the social acceptability of his behavior. A typical psychotherapist today feels that he is a responsible liaison agent between interpersonal reality and the patient, and tries to be flexible in the stance he takes, so that the patient may learn the maximum amount from his association with the therapist. The therapist having the latter goal is likely to want a great deal of diagnostic information about his patient as soon as possible. This will include not only information about the patient's unconscious, but also the more conscious information about his self-concept, his characteristic attitude towards other people, the ego strength he is able to bring to bear in various situations, his social stereotypes, etc. Thus, the therapist would seemingly benefit from the entire array of psychological information which could be obtained from a comprehensive battery of psychological tests.

Further, it seems an unwarranted assumption that diagnostic evaluation is of value only in connection with psychotherapy planning. Certainly many times important administrative decisions can be aided by the presence of such data, clinical diagnoses with many behavioral implications can be made more accurately, and research can be carried on which will enable psychologists to show more understanding and utilize their time more productively as a result. There seems to be no necessary inconsistency between spending time in diagnostic

evaluation or in research since, as was once pointed out to me by one colleague, "I am amassing research data every time I sit down to evaluate a patient."

3. "Diagnostic evaluation is so time consuming and lengthy and clumsy that it engages the psychologist far beyond the point of diminishing returns."

There is some justification in this criticism if one takes as a framework the typical activity of the psychologist of 15 years ago in carrying on diagnostic evaluation. I am sure that many readers will recall the time when it was not unusual to spend days, if not weeks, on administering, scoring and interpreting a group of projective tests and trying to gain every little nuance of "psychodynamic significance." However, this is not current psychological practice. In an effort to meet the needs of a situation as described above, the typical kind of evaluation carried on today relies in large part upon self-administered instruments such as the MMPI, the Sentence Completion Test, various adjective check lists, self-administering TAT's, and other such instruments. Typically, some of this information will be scored, tabulated and graphed by clerks and the time of the clinical psychologist is restricted to administering tests like the Rorschach, Bender and WAIS, and interpreting these together with others as a whole battery. In reading answers to questions on the ABEPP examination, the present author has discovered that the practice described above is quite typical of ABEPP candidates. They seem to feel that it gives them maximum information, and is still parsimonious as to time. This is in line with the current practice of emphasizing multi-level assessment of the individual, including evaluation of his public image, his conscious self-concept, and his underlying motivational system. Sticking strictly to projective or to intellectual instruments without integrating these with anamnestic data, interview data and the results of

objective instruments of appraisal would seem to be frivolous in terms of the time consumed, and less accurate and important in terms of the information desired.

4. "There is little point in engaging in diagnostic evaluation because of the unknown and questionable validity of the various instruments involved."

This argument, of course, can always be brought to the fore, and contains some justice. Some clinical psychologists engaged in diagnostic evaluation tend to over-evaluate the accuracy of their instruments, and to give a false impression of finality to the recipients of their reports. However, it is possible to hierarchically arrange the tests and the interpretations from the tests in terms of the degree of evidence available for validity. Certainly one can make predictions with greater accuracy from the MMPI than from the DAP. At least the probability of the truth of the prediction can be specified in quantitative terms. Further, in comparing our science with other sciences that deal with probabilities, we need feel no shame regarding the status of diagnostic evaluation methods. For example, the medical science of obstetrics is not only unsuccessfully attempting to explain spontaneous abortions in many instances, but is still trying to get a rationale for the usual nine-month gestation period. Certainly, social sciences such as anthropology and sociology are often forced to fly by the seat of their pants to at least as great an extent as our science of clinical prediction. Further, the critics of diagnostic evaluation on the lack of validity basis often are enthusiastic psychotherapists. To claim that there is more evidence for the efficacy of psychotherapy than for the predictive accuracy of psychological tests is to be blind indeed.

5. "Diagnostic evaluation has no demonstrated value in helping to make administrative decisions, and is consequently a waste of time."

This criticism hurts because it is so often justified! All too often the clinical psychologist in submitting the report of his findings talks within a self-contained frame of reference of his own, and completely fails to be responsive to the needs of the referrer for specific recommendations. When the question involves hospitalization or outpatient treatment, promotion or non-promotion, execution or non-execution of a criminal, and other similarly significant decisions, it is of very little value to have the psychological report refer to the patient's latent homosexuality, masturbatory anxiety, confusion in sexual role, difficulty in interpersonal relationships, and other similar "psychodynamic" stereotypes that are so dear to the heart of many clinical diagnosticians.

However, it is the belief of the present writer as documented in his recent book (Klopfers, 1960), that clinical psychologists can do better than this and can respond directly to the referral problems in a clear concise way. Thus, the present writer would agree that this criticism is justified, but not that it is insurmountable.

6. "Psychological reports are of no value since they are totally incomprehensible to anyone other than a psychologist."

This criticism is the unkindest cut of all, but unfortunately is justified all too often. Many of the recipients of our information find themselves alienated by the esoteric jargon and the flights of fancy so dear to the heart of clinical psychologists. They somehow sense that this is an exercise in mental masturbation rather than an attempt to communicate significant information that will have practical value. Whether our incomprehensibility is due to lack of real understanding of the concepts involved or whether it constitutes an attempt to maintain psychological distance between us and our referrers is difficult to say. At any rate, as long as such incomprehensibility continues to be manifest, it will serve as a barrier to the full accep-

tance and utilization of diagnostic evaluation by clinical psychologists.

CONCLUSIONS

Human behavior continues to be as complex as ever. There is no particular evidence of the increased sophistication of clinical psychologists in finding short-cuts to the alleviation of human suffering which will obviate the necessity of thorough idiographic evaluation. This article has expressed some doubts as to the wisdom and justification of by-passing diagnostic evaluation in clinical practice. Certainly the use of psychological tests continues to be an essential element in psychological research. A trend has been noted which consists of the greater integration of objective and real-life methods of diagnostic evaluation with those that are more projective and indirect. This trend leads to greater parsimony in the use of the psychologists' time

and more detailed information available to the referrer, leading to better decisions and less waste of human resources. Finally, the present author has stated his position that the clinical psychologist uses his resources both as a scientist and as a clinician to the greatest extent in his function as a diagnostician. There is no aspect of the job of clinical psychologist which is more complex, more difficult, and more challenging.

REFERENCES

- Guertin, Witt, et al. Research with the Wechsler Intelligence Scale for Adults, 1955-60. *Psych. Bull.*, 1962, 59, 1-27.
- Klopfer, W. G. *The Psychological Report: Use and Communication of Psychological Findings*. New York, Grune & Stratton, 1960.
- Masling, J. The influence of situational and interpersonal variables in projective testing. *Psych. Bull.*, 1960, 57, 65-85.

Received March 15, 1962

Human Movement Responses and Verbal Expression in the Rorschach Test¹

MURRAY LEVINE AND GEORGE SPIVACK

Devereux Foundation Institute for Research & Training, Devon, Pennsylvania

Rorschach's (1942) emphasis on the human movement response (*M*) has been held by some authorities (Piotrowski, 1960) to be one of his most original contributions to the development of the Rorschach test, and it certainly has been most fruitful from the point of view of research. In one viewpoint, the *M* response is taken as an index of fantasy activity. *M* is thought to be related to the ability to inhibit impulse expression and to the ability to delay gratification on the grounds that fantasy permits partial discharge of impulse through "trial action" (Rapaport, 1951). Those who give *M* responses are capable of delaying impulse expression because such individuals may resort to fantasy as a means of controlling the impulsion toward immediate expression. While there is a great deal of confirmatory evidence in adults (Singer, 1960), some studies (Neel, 1960; Fager, 1960) have been unable to confirm previous findings and the hypothesis does not seem to hold for children (Litman, 1957; Spivack, Levine, Fuschillo, Tavernier, 1959). A great deal of evidence has also accumulated showing a low, but statistically significant relationship between *M* and intelligence (Levine, Spivack, Wight, 1959), and inhibition ability has been related to intelligence (Levine, Spivack, Fuschillo, Tavernier, 1959; Spivack, Levine, Sprigle, 1959).

The theory of the delay of impulse gratification is nowhere clear in indicating why the *M* response alone should be the critical element. If it is a matter of fantasies of action, why

not consider *FM* and *m*, or other aspects of the content? Sarason (1954) believes the *M* response involves the projection of subjective or internal factors, and he suggests these same factors may be reflected in any response, whether or not it involves movement. Sarason goes on to say: "A very fruitful and important research area would be the development of a scale or continuum whereby all responses could be evaluated in terms of the degree of subjective and personal factors which they reflect" (1954, p. 220).

For some time, we have been developing what we call an "index of repression" for the Rorschach test. This index is based upon details of the language used in the response, and it may be applied to all responses. In developing this "index of repression", we were interested in relating the scores to the usual scoring features of the Rorschach test. In several samples, using both individual and group records, the highest correlation consistently has been between the index of repression and the usually scored *M* response. Other correlations appear, but the relationship to *M* has always been highest. At first glance, the correlations are not surprising since the use of verbs enters into the index of repression score, but the existence of very high correlations led us to explore whether or not there was a relationship between *M* and the index of repression, when the index was computed *independently* of the use of verbs. This study reports on the relationship between the *M* response as normally scored, the spontaneous use of verbs in the free association part of the Rorschach, and

¹This paper has benefited from criticism of an earlier version by Robert G. Ballard and Zygmunt Piotrowski, neither one necessarily agreeing with the interpretations it contains.

the index of repression computed independently of the use of verbs.

METHOD

In developing the index of repression, we have been aided by the cooperation of other workers² who made Rorschach records available to us for study. We have thus been able to accumulate large numbers of records, and have been able to test various hypotheses on several samples. We have worked with a variety of "haphazardly" chosen samples, with the assumption that consistent results would permit broad generalization of findings. In this study, we have used six different samples:

Subjects

1. Normal Adolescents: A group of twenty records of normal adolescents were collected as part of a validation study of the index of repression. The Ss were all male, and were tested at a local, suburban high school. They were assigned to the study by the guidance counselor of the school. The records were individually administered, using standard instructions and using the Rorschach blots.

2. College Females: Holtzman and Gardner (1959) reported a study of the Rorschach records of extreme levelers and extreme sharpeners. The 20 Ss were female university students, ranging in age from 18-21. The Ss were selected from a larger group of 80 on the basis of their performances in the Schematizing Test. The Rorschachs were individually administered using standard blots and instructions.

3. Student Nurses: In the standardization of the group Rorschach test, Harrower-Erickson and Steiner (1945) collected individual Rorschach records from 68 student nurses

who were also administered the group Rorschach. Half the Ss received the group test first, and half received the individual test first. Retest took place about a week later. Standard blots and standard instructions were used. The data reported here are based upon the individual records.

4. College Females: These records were collected by Page (1957) as part of a study of day dreamers. One-half of the 35 records were given by Ss who admitted to a great deal of daydreaming, while the other half of the records were given by those who were low on Page's daydreaming scale. The records were administered in group following Harrower-Erickson and Steiner's procedure.

5. College Females, Holtzman inkblots: These 44 records were collected by Young (1959) as part of a study of Witkin's field dependence hypothesis. The group version of the Holtzman inkblot test (1961) was administered following Holtzman's procedure. The Ss were female undergraduates attending a summer session.

6. College Males, Holtzman inkblots: These records were also collected by Young (1959). The Ss were male undergraduates attending a summer session.

Procedure

The index of repression is derived from a scoring of the free associations to the ink-blots according to a system devised by Levine and Spivack (1960a). Each response is scored according to the degree to which it is specific, elaborated, incorporates movement, and contains impulse content or self references.

Specific means the extent to which the content of a response refers to a unique entity, in contrast to a broadly general or vague concept. The response "animal" would reflect more repression than "bear", while "Donald Duck" would reflect even less repression, it being a unique reference. Elaboration refers to the adjectives and adverbs which expand and add

²We offer our sincere appreciation to Riley W. Gardner, Molly Harrower, Horace A. Page, Jack Quackenbush, and Harl H. Young, all of whom helped by making Rorschach records available to us, one way or another.

color to the percept reported. A response such as "tree" is taken to suggest more repressive tendency than a response such as "a dying tree with drooping branches". Consistent with a view that movement responses indicate mental activity, all movement is taken to suggest a lessened degree of repression. In general, verbs are scored in this category. Any direct reference to sexuality, hostility, anal-ity, or dependency is taken to signify a lesser reliance on repression, since it is assumed that the repression is maintained to keep just such verbal derivatives from consciousness. All statements made by the *S* during the free association regarding personal feelings and past memories are regarded as indicating a lessened degree of repression. The system allows us to weight each response in the free association part of testing, the total score for any single response being the sum of weighted points for the response. The mean of these weighted scores represents the extent to which *S* characteristically relies upon a repressive form of defense. The obtained mean score actually is an inverse measure of the amount of repressive style typical of the individual in question.

Scorer reliability has been found to be above .90 in several tests, and retest reliability of the score after one week has been shown to be .91. In children, retest reliability coefficients for periods up to ten years have been shown to average .60 (Levine, & Spivack, 1960b).

Because we were interested in the

relationship of the index of repression to *M*, and verbs are part of the index of repression score, the index of repression was computed for each *S* omitting any scoring for verbs. This includes human, animal and inanimate movement. At the same time, count was kept of the number of scorable verbs used spontaneously by the *S* (see manual for rules for scoring verbs). The *M*s were scored by those who took the original records.

RESULTS

Table I summarizes the relationship between *M* as usually scored, and the index of repression score computed without verbs. In general, the results are consistent in showing a positive relationship between the index of repression score and the *M* response, as usually scored. The correlations using the Holtzman blots may be better because of the greater reliability of scores based upon the written output to 40 cards.

The index of repression is a mean, and it controls the factor of response total. Our previous work has shown that *R* is uncorrelated with the repression index. The significant correlations between *M* and the index of repression are not based upon any common dependence of both scores on *R*. However, as a further check, a contingency coefficient was obtained between *M*% and the index of repression score computed without verbs, using Harrower's sample. The relationship was significant, and the contingency coefficient using *M* was ex-

TABLE I—The Correlation Between *M* as Usually Scored and the Index of Repression Score Computed Without Considering Verbs

Sample	<i>N</i>	<i>r</i>	<i>p</i>
Individual Rorschachs			
Male adolescents	20	.59	<.01
College females (Holzman and Gardner, 1959)	20	.41	.08
Female student nurses (Harrower-Erickson and Steiner, 1945)	68	.28	.04
Group Rorschachs			
College females (Page, 1957)	35	.21	NS
College females, Holtzman inkblots (Young, 1959)	44	.63	<.01
College males, Holtzman inkblots (Young, 1959)	48	.76	<.01

TABLE II—The Correlation Between the Total Verb Score and the Index of Repression Score Computed Without Considering Verbs

Sample	N	r	p
Individual Rorschachs			
Male adolescents	20	.59	<.01
College females (Holzman and Gardner, Female student nurses (Harrover-Erickson and Steiner)	20	.63	<.01
Steiner)	68	.35	<.01
Group Rorschachs			
College females (Page)	35	.43	.02
College females, Holzman inkblots (Young)	44	.69	<.01
College males, Holzman inkblots (Young)	48	.62	<.01

actly the same as the coefficient using $M\%$, .38 in both cases.

Table II summarizes the correlations between the total number of verbs scored and the index of repression score computed without reference to verbs. The correlations are uniformly significant, and fall within a relatively (for Rorschach work) narrow range. It is clear that Ss who produce more verbs also produce richer responses in other respects. Ss who produce many verbs are more likely to embellish their responses with the other elements scored (adjectives; specificity in the central noun of the response; more likely to produce responses containing direct references to impulses).

Our evidence does not support the hypothesis that there is a general verbosity which is characteristic of Ss. Page (1956) also administered TATs to his sample, and we found no relationship between the total number of words S produced on the TAT, and the full index of repression score.

The TATs had previously been rated by Page for "level of interpretation", on an eight point scale ranging from "rejection" and "description of physical appearance" to "complete narrative with interpretation of feelings, attitudes and/or thoughts of characters." A score for each S was based upon the mean rating for six TAT cards. The correlation between the full index of repression score and the mean "level of interpretation" rating was .35, significant at $p = .05$ for $N = 35$. However, the correlation between M alone and the TAT "level

of interpretation" was .10, a value which is not significant.³ Apparently the additional features of the index of repression add a significant component to the evaluation of Rorschach responses.

DISCUSSION

The same psychological tendency that leads an S to enrich his response with movement is probably found in his tendency to enrich his response with adjectives and adverbs, to be more specific in his response (e.g. "George Washington" versus "a head"), and to permit impulse material into his responses. Our data suggest the M response is not unique in itself, but rather it is part of the tendency to produce well defined, elaborated and individualized responses, which tendency is quantified for all responses by the index of repression measure.

The viewpoint developed here suggests that the Rorschach method elicits a characteristic style of verbal expression, a style which is reliable and which is established very early in life (Levine and Spivack, 1960b). While it is possible to interpret the content of movement responses in

³The level of interpretation rating used by Page seems to have much in common with Weisskopf's transcendence index (1950) and the transcendence index has been shown to be related to the M response by Singer and Herman (1954) and by Herman (1957). While Page's rating scale shares much in common with Weisskopf's procedure, the two are not identical. That no significant correlation was obtained between M and Page's rating scale scores could very well be due to differences between his scale and the transcendence procedure.

many other ways, as Piotrowski (1960) does, our data suggest that movement responses, which must, of course, be expressed as verbs, are part of a style of rich verbal expression. It is possible to observe this style of expression in responses other than those relatively few which happen to contain verbs.

The correlation between the TAT level of interpretation rating and the index of repression suggests the tendency to give rich and elaborate verbal fantasy productions is a general psychological characteristic. Sarason's (1954) opinion that subjective and internal factors may be evaluated just as well in any response, and not in *M* responses alone, receives some support from these findings.

If we may deal with a tendency to use rich and specific language, why should this tendency show a relationship to the ability to delay impulse expression (assuming the index of repression score would have the same correlates as the *M* response)? Murphy (1947), writes that personality organization changes from the global and the diffuse to the specific and the integrated and that as far as the perception of self and others is concerned, the increasing differentiation is accompanied by, if not supported by, an increasing use of more specific names for parts, for people, for objects and for feelings. Osgood (1953) and William James (1950) long before him, have pointed out the utility increasingly varied language may have in mediating ever finer and more complex discriminations, and both Dollard and Miller (1950) and Rogers (1951) have emphasized the role of labeling in integrating experiences and sensations into the self. It would appear that ready availability of a variety of language forms would reflect increased possibilities for relating to the self and to one's environment. When there is more limited language availability, flight, or less refined and developmentally more primitive means of expression may be

the primary alternatives. Greater variety and availability of language may be accompanied by a greater differentiation of personality organization, more complex layering of defenses, and the language may mediate more complicated sequences of responses.⁴

In the field of psychopathology, with greater language availability, one might expect more deliberate behavior, more specific symptom formation, subtle guilt reactions, depression, and less tendency toward the unmodulated, diffuse type of discharge of tension as in hyperactivity, hyperaggressivity, temper tantrums and poor frustration tolerance, or in developmentally primitive denial, or flight.

The concept of the delay of discharge of tension does not necessarily imply that tension is never discharged, and that acting out never occurs in individuals who employ rich and expressive language. Attitudes, situations, feelings and values will determine whether or not acting out occurs at any given point in time. However, one would expect a great deal more rigidity of thought and action and a greater dependence of behavior on immediate circumstances when language to mediate more complicated sequences of responses is less available. The psychological symptom picture should be very different in individuals who have developed a quantity of differentiated language and in individuals who use language to a lesser degree.

SUMMARY

Significant correlations were shown between *M* responses as usually scored and an index of repression scored independently of the contribution of verbs, in six independent samples. The production of verbs also correlated significantly with the index of

⁴However, recent evidence has suggested the index of repression score may have different correlates in females than in males, and this hypothesis may hold only for males. (Spivack, Levine and Brenner, 1962).

repression score based upon the use of adjectives, adverbs and open impulse content. The index of repression score was shown to reflect something more than just verbosity. The index did not correlate with the number of words given in the TAT. The index of repression did correlate significantly with a rating similar to the "transcendence index" for TAT stories, but the M response alone did not correlate significantly with this TAT measure.

The results support a view of movement responses as stemming from the same psychological tendency which results in enriched responses whether or not such responses contain verbs. It is proposed that the language in which Rorschach percepts are expressed constitutes primary test data. The richer and more available the language for expression, the greater the possibilities may be for differentiated and partial reactions to one's self and to one's environment.

REFERENCES

- Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw Hill, 1950.
- Fager, R. E. Relation of Rorschach movement and color responses to cognitive inhibition. *J. consult. Psychol.*, 1960, 24, 276.
- Harrower-Erickson, Molly R., & Steiner, Matilda E. *Large scale Rorschach techniques: a manual for the Group Rorschach and Multiple Choice Test*. Springfield, Ill.: Charles C. Thomas, 1945.
- Herman, J. L. Ideational and motor correlates of the Rorschach experience type. Unpublished doctor's dissertation, New York University, 1957.
- Holtzman, W. H., Tharpe, J. S., Swartz, J. D., & E. W. *Inkblot perception and personality*. Austin: University of Texas Press, 1961.
- Holtzman, P. S., & Gardner, R. W. Leveling and repression. *J. abnorm. soc. Psychol.*, 1959, 59, 151-155.
- James, W. *Principles of Psychology*. New York: Dover Publications, 1950.
- Levine, M., Spivack, G., & Wight, B. The inhibition process, Rorschach human movement responses, and intelligence: some further data. *J. consult. Psychol.*, 1959, 23, 306-312.
- Levine, M., & Spivack, G. *The Rorschach index of repression. Manual of instructions for scoring*. Devon: The Dexeux Foundation, 1960 (a).
- Levine, M., & Spivack, G. Stability of the Rorschach index of repression. Paper read at Eastern Psychol. Assn., New York, April 1960 (b).
- Litwin, Dorothy. Sensory-tonic theory as related to the development of inhibition and social maturity. Unpublished doctor's dissertation, New York University, 1957.
- Murphy, G. *Personality*. New York: Harpers, 1947.
- Neel, Ann F. Inhibition and perception of movement on the Rorschach. *J. consult. Psychol.*, 1960, 24, 224-230.
- Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford Univ. Press, 1953.
- Page, H. A. Studies in fantasy, daydreaming and the TAT. *Amer. Psychologist*, 1956, 11, 392 (abstract).
- Page, H. A. Studies in fantasy: Daydreaming frequency and Rorschach scoring categories. *J. consult. Psychol.*, 1957, 21, 111-114.
- Piotrowski, Z. A. The movement score. In Maria A. Rickers-Ovsiankina (Ed.) *Rorschach Psychology*. New York: Wiley, 1960. Pp. 130-153.
- Rapaport, D. *Organization and pathology of thought*. New York: Columbia Univ. Press, 1951.
- Rogers, C. *Client centered therapy*. Boston: Houghton-Mifflin, 1951.
- Rorschach, H. *Psychodiagnostics: a diagnostic test based on perception*. Bern: Hans Huber, 1942.
- Sarason, S. B. *The clinical interaction: with special reference to the Rorschach*. New York: Harper, 1954.
- Singer, J. L. Delayed gratification and ego development: implications for clinical and experimental research. *J. consult. Psychol.*, 1955, 19, 259-266.
- Singer, J. L. The experience type: some behavioral correlates and theoretical implications. In Maria A. Rickers-Ovsiankina (Ed.) *Rorschach Psychology*. New York: John Wiley and Sons, 1960. Pp. 223-259.
- Singer, J. L., & Herman, J. L. Motor and fantasy correlates of Rorschach human movement responses. *J. consult. Psychol.*, 1954, 18, 325-331.
- Spivack, G., Levine, M., & Brenner, B. Sex, verbal style and the use of the semantic differential. Paper read at EPA, Atlantic City, April, 1962.
- Spivack, G., Levine, M., Fuschillo, Jean, & Tavernier, Ann. Rorschach movement responses and inhibition processes in adolescents. *J. proj. Tech.*, 1959, 23, 462-466.
- Young, H. H. A test of Witkin's field-dependence hypothesis. *J. abnorm. soc. Psychol.*, 1959, 59, 188-192.
- Weisskopf, Edith A. A transcendence index as a proposed measure in the TAT. *J. Psychol.*, 1950, 29, 379-390.

Received February 6, 1962

Post-Doctoral Training for Clinical Competence¹

MARTIN MAYMAN
The Menninger Foundation

There is still much uncertainty among students and program directors concerning the need for supplementary post-doctoral training in clinical psychology, and uncertainty too about the content and duration of an optimally useful post-doctoral program of training. But one thing seems certain: that post-doctoral training in clinical psychology is here to stay.

A brief review of some recent developments may serve as a useful introduction to my own views.

1. Ten years ago formal post-doctoral clinical training (as distinct from on-the-job training) was available at only three clinical training centers in this country. Today there are twenty-four such programs supported by the U.S. Public Health Service, and there may well be over thirty in operation by January 1963.

2. The demand for these training Fellowships has been great, and is by no means limited to newly-graduated or poorly trained clinical psychologists. For example, applications for two post-doctoral Fellowships offered in one such program have run to between fifteen and twenty each year, and there have been three times that many more inquiries. More than half these applicants had at least two years of full-time clinical experience; almost a fourth of them were two or more years beyond their doctorate; and 7 out of 10 were graduates of APA-approved clinical programs (Mayman, 1959). And, I might add parenthetically that completing one of the application forms for this program is so tedious and so time-consuming as to eliminate people only casually interested in further training.

3. The rapid proliferation of post-

doctoral training programs is paralleled by the growing number of papers which have appeared in the journals and in our convention programs concerned largely or exclusively with clinical training in general and post-doctoral training in particular. A quick survey of the literature from 1955 to the present yields the remarkable total of twenty-seven papers, twenty-six brief communications, nineteen symposia, eight published committee reports, and four published collections of papers—in all, one-hundred-fifty-seven titles on graduate and post-graduate training for clinical psychologists which have appeared in just five years.²

4. The discussion of terminal sub-doctoral clinical training all but disappeared from the literature by 1957, and was apparently dealt its death blow by the 1959 report and recommendations of the APA Committee on Sub-Doctoral Training (1959). At the same time, the number of discussions of post-doctoral training have increased from a bare total of two brief communications in 1957 to a total of nineteen titles by 1961.

5. Add to these signs of the times the repeated references made to post-doctoral training at the Miami Conference; the concern of the recently organized Corresponding Committee of Fifty with, among other things, problems of training; and the recent appointment by Division 12 of a Committee on Training Issues to guide universities, practicum programs and the E. & T. Board in their efforts to improve the scope and quality of clinical training—and we can say with assurance that “ferment” and “change” appropriately describe the state of affairs in clinical training today.

¹ Presented to the annual meeting of the American Psychological Association, Sept. 2, 1961, New York City.

² This bibliography is available on request.

A variety of post-doctoral programs have emerged during this period of vigorous search for ways and means to improve the professional preparedness of clinical psychologists. One program describes itself as offering advanced training in general clinical psychology, that is, it gives fairly advanced clinical psychologists a two year professional moratorium in which to "re-examine their role as clinicians, experiment with changes in that role, and consolidate within that role the set of skills, concepts and values which best afford them both the conviction of an identity and the impetus to further growth" (Ekstein & Mayman, 1957). This program tries to accelerate advancement from that level of competence normally reached in the early post-graduate years of clinical practice, to a level normally achieved only after many years of clinical experience.

Other clinical centers have developed programs with quite different, more highly specialized emphases. At Massachusetts General Hospital, for example, an interdisciplinary group has experimented with the training of psychological specialists in the public health field. Worcester State Hospital emphasizes the research training of clinical psychologists. Other programs have offered more limited specialization, such as, for example, the intensive training in child psychotherapy offered at the Judge Baker Foundation, and the four-year post-graduate training program in psychotherapy recently initiated at New York University.

A third broad type of post-doctoral training is the deferred internship. Post-doctoral training positions in the V. A., for example, offer first or second year internships to psychologists whose graduate training does not yet qualify them for clinical appointments.

Still another form of post-doctoral training which has become popular in recent years is the brief but intensive Post-Doctoral Institute, limited to no

more than a two-week period of morning, afternoon, and often also evening seminars or workshops on some particular aspect of clinical practice. These are effective as refresher courses or extension courses for experienced clinicians, but are of more limited value for the newly trained clinical psychologist seeking more training.

Finally, a fifth pattern for post-doctoral training deserves mention, not because of its prevalence, but because of its potential importance. That is, a program which would enable clinical instructors on leave from their universities to renew their contact with clinical problems and clinical practice.

Valuable as such heterogeneity may be at this stage in the growth of our profession, it has one significant drawback. It tends to give the term "post-doctoral training" a broad and therefore somewhat ambiguous meaning. It is important to recognize that each kind of program is uniquely suited to the need it was designed to meet. Those psychologists seeking an introductory internship which was deferred until after completion of their doctoral work should not substitute for that internship either an advanced internship or one which offers only limited specialization. On the other hand, those seeking advanced specialization can hardly substitute a series of short-term institutes or workshops for the kind of program they really need. But it is the special need which exists for the advanced internship form of training—a need which is not suitably met by available workshops, institutes, or programs offering advanced specialization—which I want to discuss at some length. To appreciate this need, one must back up a step or two and consider it in relation to a larger need.

The present ground-plan for clinical training was agreed upon at the Boulder Conference in 1948. The science and the profession of clinical psychology have changed in the almost fifteen years since then. It has

become increasingly clear that our current blueprint for graduate and post-graduate education may no longer be adequate, not only with regard to the clinical-professional training provided for, but even more important, inadequate in the presentation of those theories and research findings which will provide a sound scientific base for clinical practice.

Most clinical psychologists are familiar with Kelly's (1961) recent survey of the professional scene in clinical psychology today. One of his conclusions was that, "A fairly large number of our members (that is, members of Division 12 of the APA) do not regard clinical training as now offered as an adequate preparation for clinical practice". Forty-three percent of the opinions he sampled favored the development of a new, completely separate "professional school" for the next generation of clinical psychologists. There are those who are quick to infer from Kelly's findings confirmation for their belief that a large segment of clinical psychologists are impatient with, and would do away with, the more "tough-minded," "hard nosed," scientific training required by our universities. I suggest that precisely the opposite may be true. It is regrettable that the survey did not ask for suggestions regarding an appropriate curriculum for this new professional school for clinical psychology. Such suggestions might have been very revealing. Kelly did determine that clinicians reach most nearly unanimous accord (85%) on the need to train clinical psychologists in the science as much as in the art of clinical practice. It may be in precisely this area that clinicians find current graduate programs most deficient. Clinicians have, I believe, grown disappointed not with science but rather with "scientism," not with psychology but with too narrow a representation of psychology. They do want to learn clinical skills, but also a richer understanding of their tools.

I, for one, would prefer to accept for post-doctoral training someone who, though unable to score and interpret a Rorschach according to some favorite book of instructions, has acquired instead an appreciation of the organismic psychology of Goldstein, Werner and Stern; the developmental psychology of Piaget; Gordon Allport's and Gardner Murphy's books on personality; a working familiarity with the Gestalt psychology of Koffka and perhaps even the "act" psychology of Brentano. I would prefer that an applicant know the experimental literature on expressive movements which preceded the flowering of projective testing in this country than that he have acquired only technical competence in these diagnostic techniques. It would be well if he had more than cursory contact with the social psychology of George Mead and the social anthropology of Margaret Mead, if he were conversant with the experiments on imprinting and internal releaser mechanisms, and knew some of the new findings on brain structure and neurophysiology. No. Clinicians do not want *less* science in their basic training; they want *more* science, but a science more focally concerned with processes of life, growth and adaptation.

The four deans of the Committee on Policies in Graduate Education (1957) have said of professional training in graduate schools, that, "One can say that law and business are *professional* programs and thus can be easily defined, but that the Ph.D. is essentially an individual matter between student and master; that it is therefore filled with unpredictable elements and that, in short, it is non-professional and therefore cannot be laid out so neatly in terms of years. We agree. The Ph.D. is not a professional degree. No degree could be called professional which sets out to nurture individual discovery and which exalts newness in knowledge". Those responsible for training clinicians would probably endorse with

enthusiasm such a graduate program for the clinical psychologist, as would most clinical trainees. It may well be the misdirected nurturance of individual discovery and the smothering of their creative interest in human behavior that has led so many clinicians to feel as they do about their graduate studies.

However, progress toward the Dean's Committee training would necessarily limit a student's progress toward another ideal projected by the American Board of Examiners in Professional Psychology for the mature, clinical practitioner: that he be "a mature, integral human being, sensitive and deeply aware, highly skilled in the use of his professional tools, broadly educated—and still educable—in human affairs, artistically competent in professional practice while maintaining a scientific orientation both to the evolving knowledge in his field and to the assessment of his own performance, clearly demonstrating a keen awareness of the ethical issues involved in an intricate and unstructured professional role, and withal, joining with his fellows in scientific and professional organizations to confront general problems of common concern" (Kelley, Sanford & Clark, 1961). The Board of Professional Examiners assumes that it takes no less than five years of practice to successfully approximate this professional ideal. In that time, it is hoped, the psychologist will have learned not only all he needs to know for sound practice, but also how to think and feel and use himself as a clinician.

As university graduate departments more and more eschew the responsibility for looking after the student's progress toward this professional-practitioner ideal, most of this responsibility is left with the individual himself, often after no more than one year of supervised experience. The consequences of such a course are not always salutary. I am reminded of the disaffection which Kelly and Goldberg (1959) found so widespread

among clinical psychologists ten years after they entered clinical training. A shockingly high number of them—two out of five with Ph.D.'s and three out of four with terminal M.A.'s—no longer believed in what they were doing and would choose some vocation other than clinical psychology if they could live their lives over again. Perhaps they no longer believed in what they were doing because they never learned to do it well. Abortive training is certain to undermine the clinician's sense of personal integrity. In the long run it will also surely undermine the integrity and progress of clinical psychology itself.

Clinical psychology today seems to be confronted squarely with this dilemma: Graduate programs cannot be expected to provide professional training when it is a full time job providing the kind of academic training which lays a solid foundation for later professional work. Internship training provides a barely adequate start toward the development of professional effectiveness. The number of programs available which offer advanced training is not commensurate with the size of the problem. Nor is there even an adequate blueprint available which can guide the newly arrived clinician toward full professional competence.

There are, however, some things which can be done now to attack this problem. First, a survey should be undertaken to determine what level of competence is actually achieved by clinical psychologists upon completion of their doctoral studies. This survey should sample the opinions of students, program directors and prospective employers and supervisors of the post-graduate clinician. Secondly, a study group should be appointed to consider how the profession may best help young clinicians fill the gaps in their training and accelerate their development into sound, skillful clinicians.

REFERENCES

American Psychological Association, Educa-

- tion and Training Board Task Committee; Report on the role of the master's degree in doctoral training. *Amer. Psychol.*, 1959, 14, 501-503.
- Association of Graduate Schools, Committee on Policies in Graduate Education; Doctor's and Master's—Good and Bad! New York City meeting, October 22, 1957.
- Ekstein, R. & Mayman, M.; On the professional identity of the clinical psychologist, *Bull. Menn. Clin.*, 1957, 21, 59-61.
- Kelley, N. H., Sanford, F. H., & Clark, K.; The meaning of the ABEPP diploma, *Amer. Psychol.*, 1961, 16, 132-141.
- Kelly, E. L.; Clinical psychology—1960: Report of survey findings, *A.P.A., Div. 12 Newsletter*, Winter, 1961, 1-11.
- Kelly, E. L., & Goldberg, L. R.; Correlates of later performance and specialization in psychology, *Psychol. Monogr.*, 1959, 73, No. 482.
- Mayman, M.; Five-year review and prospectus: The Menninger Foundation Post-Doctoral Training Program in Clinical Psychology, 1959, unpub., 12 pp.

Received March 16, 1962

Effects of Training on Students' Test Interpretations^{1,2,3}

RUDOLF H. MOOS

University of California School of Medicine and
The Langley Porter Neuropsychiatric Institute

This study attempts to deal with one aspect of the question of how students learn clinical skills. It attempts to assess the effect of one graduate course in diagnostic methods on students' Q-sort interpretations of two commonly used tests.

Rodgers (1957) was concerned with the sources of information that enter into an interpretation of a test, specifically the Rorschach. He assessed the kinds and degrees of influence on blind Q-sort interpretations made by students in a beginning course on the Rorschach, and found that, even though instruction seemed to significantly influence the interpretation of the protocols, "it accounted for only three to four per cent of the probable raw score variance, while about 58 per cent of the variance had to be considered error variance." He did not use a control group to determine the effect of mere repetition of the task, and pointed out that the significant change could possibly be "due to repetition of the task alone, with the didactic material being a total waste except as interposed material to

fill time between trials one and two." Rodgers also did not deal with the question of whether the Q-sort interpretations were valid or not, since the study was concerned exclusively with the sources of variance, whether or not this variance contributed to valid interpretations.

The present study, in addition to replicating Rodgers' work with the Rorschach, had three new aspects. 1) The results of training in a second test, the Minnesota Multiphasic Personality Inventory (MMPI), were compared with the results of training in the Rorschach. 2) The change in validity of Q-sort interpretations was investigated by having three experienced clinicians criterion sort protocols of each test. 3) The effects of mere repetition of the task were assessed by the use of control groups.

METHOD

Subjects

There were four groups of subjects, two experimental groups and two control groups. The subjects were all graduate students in clinical psychology. During the first semester of their second graduate year students in the clinical sequence receive beginning instruction in seminars devoted either to the Rorschach or to the MMPI. The two experimental groups consisted of all students in the Rorschach ($N = 4$) and MMPI ($N = 9$) seminars. During the first graduate year students in the clinical sequence receive beginning instruction in seminars devoted to intelligence testing and beginning casework with children. The two control groups ($N = 6$ per group) consisted of the students in this seminar. They, of course, received no formal training in either the Rorschach or the MMPI.

¹The writer would like to express his appreciation to David A. Rodgers, under whose supervision this research was conducted, and to Guy Hamilton Crook, Harrison Gough, Frank Harper, Betty Kalis, William O'Donnell, Virginia Patterson, David Rodgers, John Starkweather, and Richard Wolton, all of whom served as criterion sorters.

²This research was supported in part by funds made available from USPHS Grant #2M-7082 NJ and was conducted during the term of a USPHS fellowship at the University of California School of Medicine and the Langley Porter Neuropsychiatric Institute. The author's present address is, Dept. of Psychiatry, Stanford University School of Medicine, Palo Alto, California.

³This paper was read, in part, at the annual meeting of the American Psychological Association, New York City, September 1, 1961.

Protocols⁴

The MMPI and Rorschach protocols of two inpatients of a Veterans Administration general medical and surgical hospital were used. These male patients were in a small therapeutic community and were diagnosed as dissociative reaction (CM) and depressive reaction (JA). The protocols were selected for representativeness and because the author had had extensive individual experience with both patients. The MMPI profile was dittoed, the profile including only the three validation and the nine clinical scores. The Rorschach protocols were dittoed verbatim typescripts and included a summary of scoring determinants and locations on an additional page. There was no identifying information on the protocols except the age and sex of the patients, and no additional information was supplied during the study.

Q-decks⁴

Two different Q-decks were used, one being designated the MMPI deck and the other the Rorschach deck. The decks were so designated because of the manner in which they were constructed. The MMPI Q-deck was a 154 item deck constructed by Halbower (1955), who obtained the items from psychological reports, intake summaries, case histories in the MMPI ATLAS, and several other sources. The items appeared to be particularly relevant to the kind of information that can be fruitfully abstracted from an MMPI protocol, and they had been used in two studies assessing clinicians' MMPI interpretations (Halbower, 1955; Enright, 1959).

Some examples of items in this deck are (a) "presents himself as being physically or organically sick," (b) "tends to ignore, gloss over, and rationalize away problems and areas of conflict with the 'pollyanna' attitude that things will probably turn

out for the best," (c) "is defensive about admitting psychological conflicts (tries to avoid revealing himself as having psychological conflicts and emotional distresses)." The items were sorted into a forced choice distribution, from most characteristic to least characteristic, with the following number of items in the piles: 2, 6, 10, 18, 25, 32, 25, 18, 10, 6, 2.

The Rorschach Q-deck was developed by the author in the following manner. All interpretive statements in the interpretation section of Klopfer's (Klopfer, *et al*, 1954) volume on Rorschach technique were listed. These included both the major interpretations (presented in bold faced type in the book) and any other statement which was presented as an interpretation for an individual Rorschach. This procedure resulted in approximately 300 statements, which were reduced to 108 statements, the major criterion of exclusion being to reduce redundancy. These 108 items constituted the Rorschach Q-deck. Some examples of items in this sort are: (a) "draws heavily upon own needs to restructure perceptual world," (b) "has long range goals in terms of which he can deny immediate satisfactions without feeling too much frustration," and (c) "preoccupied with pregenital sexuality." The items were sorted into a forced choice distribution, from most characteristic to least characteristic, with the following number of items in the piles: 1, 6, 18, 29, 29, 18, 6, 1.

The essential differences between the two decks was that the MMPI deck dealt with slightly more surface, objective, indices of behavior, such as symptoms, while the Rorschach deck dealt with behavior somewhat less subject to empirical validation. The Q-decks appeared to reflect the differences in the kinds of interpretations made with Rorschach and MMPI protocols.

Procedure

At the first meeting of the MMPI

⁴The author will, upon request, make available copies of the protocols and the Q-decks.

The Rorschach criterion intercorrelations are generally lower than those with the MMPI. As in the case of the MMPI, JA's Rorschach protocol elicits more agreement than does CM's. The two Q-decks appear to make some difference here, that is, the Rorschach deck elicits more agreement among the sorters than does the MMPI Deck especially with CM's Rorschach protocol. In general, however, experienced sorters appear to do about equally well with the two decks, regardless of which protocol they happen to be sorting. It should be noted that there is no external criterion of validity; this part of the study is concerned only with the inter-sorter reliability of the criterion sorters.

All criterion correlations were intercorrelated, and the average intercorrelations are presented in Table II. The results indicate that the two Rorschach protocols are sorted quite differently, as are the two MMPI protocols. That is, the results cannot be explained by assuming sorting on the basis of a dimension of general maladjustment. The two Rorschachs and the two MMPI's are seen as having a great deal of unique characteristics; most of the variance is specific to the particular protocol.

If we assume that both tests are measuring the same underlying dynamics, and that these dynamics are reflected in both tests, then there should be high average intercorrelations between interpretations made from the same patients' different tests. The average intercorrelations between CM's Rorschach and MMPI protocol interpretations, and between JA's Rorschach and MMPI protocol

interpretations, indicate that this is not true, since neither of them accounts for more than about 7% of the total variance. In fact, it does not matter whether the test protocols are from the same or from different people. The average intercorrelations between the Rorschach and MMPI of the same patient are no higher than those between the Rorschach of one patient and the MMPI of the other. It is, of course, not possible to say whether the pattern of intercorrelations are specific to the particular patients and/or protocols selected; however, it should be noted that neither was pre-selected in order to achieve any particular desired pattern.

In order to discover how effective instruction had been in raising student-criterion correlations, a separate analysis was conducted for each test. Since it was felt that the seminar instructors might have a great influence over the kinds of interpretations made by the students, two separate analyses for each test were made. The first utilized only the student-instructor correlations while the second utilized all three student-criterion correlations.

A difference score was first obtained for each sort (the after instruction minus the before instruction correlation with the instructors' sort), and these differences were then tested by a simple *t* test. Next the three before instruction student-criterion sorts were averaged, as were the three after instruction student-criterion sorts, and new difference scores were obtained. The resulting differences were tested by new, but of course not independent, *t* tests.

There were no significant differ-

TABLE II—Average Criterion Q-sort Intercorrelations

1 Patient-Protocol	2 Patient-Protocol	Rorschach Q-deck	MMPI Q-deck
CM-ROR	CM-MMPI	.26	.25
JA-ROR	JA-MMPI	.13	.17
JA-ROR	CM-ROR	.01	.06
JA-MMPI	CM-MMPI	.26	.09
JA-MMPI	CM-ROR	.26	.25
JA-ROR	CM-MMPI	.14	.16

ences between before and after instruction student-criterion correlations in the Rorschach experimental group, regardless of whether only the instructors' sort or the average of three sorts was used as the criterion. For the MMPI experimental group the differences were significant (p less than .01) in both cases.

The control group data were analyzed in exactly the same manner as the data in the two experimental groups. None of the four analyses (Rorschach control group, MMPI control group, each using both the student-instructor correlation and the average of the three student-criterion correlations), showed a significant difference when the before "instruction" student-criterion correlations were compared with those after "instruction."

These results appear to show that instruction was effective in the MMPI class while it was not effective in the Rorschach class, and that the mere repetition of the task, as occurred in the control groups, had no effect.

When the control groups were compared with the experimental groups, however, there were no significant differences in the changes shown in the before-after student-criterion correlations. The MMPI control group showed a slight increase from before to after "instruction" student-criterion correlations, and it is because of this increase that the MMPI experimental group did not show a significantly greater increase than its control. It is notable that the Rorschach control group changed slightly in the direction of higher student-criterion agreement without instruction, while the experimental group, with instruction, showed very slightly lower student-criterion agreement.

The fact that there were no significant differences between the changes in the experimental and control groups raises the possibility that the changes which did occur were due simply to the students' increased so-

phistication in personality dynamics. It may be that increased agreement on MMPI protocols depends more on general clinical sophistication than on specific training in the MMPI itself. It appears plausible that, because the MMPI scales are in terms of diagnostic or symptom categories, they might reflect the interpretation "more hysteria and related mechanisms," rather than the interpretation "more elevation on scale 3 of the MMPI." These considerations are not very encouraging for formal didactic classes in either MMPI or Rorschach interpretive theory.

When the changes with instruction for the two tests were compared to each other, again using difference scores and t tests, the results indicated that the changes in the MMPI group were significantly greater than the changes in the Rorschach group. The difference was significant (p less than .05) regardless of whether the student-instructor criterion correlation alone was used, or whether all three student-criterion correlations were averaged. It should be noted again that the comparison is not between instruction in the MMPI and instruction in the Rorschach *per se*, since the MMPI students received some instruction in the method of Q-sorting. It is not possible, in this study, to separate the effects of instruction in the tests and instruction in the Q-sort method. It is, however, still a significant finding that the Rorschach sorts showed essentially no change in their student-criterion correlations with instruction.

It should be noted that the actual administration of a Rorschach is usually considered an important part of Rorschach testing and that the blind interpretation of Rorschach protocols is not standard clinic procedure. This study does not, therefore, assess the normal interpretive procedure used with the Rorschach. It is possible that the three criterion sorts would have been intercorrelated to a much higher degree if the three sorters had

each independently tested the two patients. It is also possible that independent student testing might have raised the correlations. In contrast, the "blind" interpretation of MMPI profiles is a clinically conventional procedure, and therefore appears to be an adequate assessment of the normal handling of the test. The clinical validity of the Rorschach may be due to its use as a structured interview by experienced clinicians, rather than to the validity of the classical interpretive hypotheses.

Separate analyses, done in the same manner as described above, were conducted to discover whether there were any differences between the two Q-decks or between the two patients' protocols in the extent of the changes they showed with instruction. None of the differences were significant. These results tend to confirm the hypothesis that there are no differences between the patients' protocols and to affirm the hypothesis that the two Q-decks reflect instructional changes differentially. It appears that student and criterion sorters, whether sorting the Rorschach or the MMPI, agree as much on placements of dynamic or "deep" traits as they do on the placements of symptom oriented and relatively more observable patient characteristics. These results weigh against the accepted idea that the Rorschach

is more effective for making depth interpretations about personality while the MMPI is more effective for making symptom oriented statements. Actually the MMPI appears to be more effective in both. It should be noted that "effective" means that a high correlation can be achieved between students and expert clinicians, not necessarily that expert clinicians can agree with external (non-test) criteria. The study presents no evidence on the latter point.

The before instruction student-criterion Q correlations were averaged, for each of the four groups, in order to see whether the experimental and control groups were comparable at the outset. These figures, as well as the average after instruction student-criterion correlations, are presented in Table III.

The results indicate that the experimental and control groups were roughly comparable before instruction. The average correlation for the MMPI control group increased without instruction, so that by the end of the semester they were doing as well as, or even slightly better than, the MMPI experimental group was before instruction. This is reasonable if we remember that the students in the control group will be taking the MMPI seminar the following year. The control group increase lends

TABLE III—Average Student-Criterion Correlations

Groups	Student-Instructor Correlations		Average of 3 Student- Criterion Correlations	
	Before Instruction	After Instruction	Before Instruction	After Instruction
Exp. MMPI N = 31	.32	.43	.32	.42
Control MMPI N = 9	.27	.32	.31	.36
Exp. Rorschach N = 15	.17	.16	.16	.15
Control Rorschach N = 9	.18	.19	.18	.24

N = # of Q-sorts averaged (odd numbers result because some subjects completed only one sort.)

some support to the notion that learning about general personality dynamics, apart from learning about a particular test, may slightly improve the accuracy of interpretations about that test.

The results further tend to cast doubt on the often made hypothesis that information contained in a Rorschach protocol gives naive subjects much more "face valid" information from which interpretations about the testee may be made than do just the scores on an MMPI profile. The naive sorters actually do better with just the MMPI than they do with all the information contained in the Rorschach protocol. The added information does not result in added valid variance in the student interpretations.

In order to obtain a rough measure of the average agreement of the criterion sorters all the Rorschach and MMPI criterion intercorrelations were averaged. The average criterion intercorrelations were .52 for the MMPI and .31 for the Rorschach. The average agreement between students and criterion sorters, after only one semester of instruction in the MMPI, approaches the agreement that the criterion sorters are able to achieve among themselves. This conclusion does not appear to be as true of the Rorschach although the average student-criterion correlations are only slightly more discrepant from the average criterion intercorrelations than in the case of the MMPI.

The amount of variance accounted for by these intercorrelations is quite low. The criterion sorters for the MMPI, those showing the highest intercorrelations as a group, show common variance ranging from about 10% to no higher than about 50%. The average student-criterion correlations in all cases represent less than 20% of the variance. Over 80% of all variance in all tests must then be regarded as error variance, a fact which is not very encouraging for either the MMPI or the Rorschach.

This indicates that the superiority of MMPI instruction over Rorschach instruction, or student interpretations based on the MMPI over student interpretations based on a blind Rorschach, represents the ability to reduce error from percentages in the 90's to percentages in the 80's.

SUMMARY

This study into the changes of Q-sort interpretations with instruction has shown that a semester's exposure to the MMPI and to the method of Q-sorting results in a significant increase in the degree of agreement between student and expert Q-sort interpretations of this test, while a semester's exposure to the Rorschach results in no change. Groups who have no specific exposure to either test or to the method of Q-sorting also show no changes. The fact that the MMPI seminar included some instruction in the method of interpretation utilized, i.e., Q-sorting, and the Rorschach class did not, confounded the direct comparison between the effects of test instruction in the two seminars, although the MMPI group did improve significantly more than did the Rorschach group. There were no differences between interpretations made with a Rorschach Q-deck and an MMPI Q-deck. Inter-sorter agreement with the MMPI was consistently higher than with the Rorschach, regardless of which patient, which Q-deck, or which subject group was used.

These findings, in conjunction with Rodgers' (1957) earlier findings, cast some doubt on the value and result of formal instruction in the Rorschach. Rodgers found that only about 4% of the probable raw score variance was related to instruction, and in this study there was no change whatever toward greater agreement with criterion sorters after instruction. No conclusions may be reached from these results as to whether the Rorschach can be profitably and validly used in a clinic setting; however, it

may be that beginning instruction in the Rorschach method in field placements and in individual supervisory sessions would be sufficient to learn the technique. The findings clearly emphasize the large amount of error variance in test interpretations, and support the often cited superiority of the inter-sorter agreement which can be achieved in the interpretations of empirical objective tests, such as the MMPI, in which the rules of interpretation are made explicit and are relatively easily communicated.

REFERENCES

Enright, J. B. Profile types and prediction

from the MMPI. Unpublished Ph.D. dissertation, University of California, Berkeley, 1959.

Halbower, C. C. A comparison of actuarial vs. clinical predictions to classes discriminated by the MMPI. Unpublished Ph.D. dissertation, University of Minnesota, 1955.

Klopfer, B., Ainsworth, M. D., Klopfer, W. G., Holt, R. R. *Developments in the Rorschach Technique*. Volume I. New York: World Book, 1954.

Rodgers, D. A. Sources of variance in students' Rorschach interpretations. *J. prof. Tech.*, 21, 1957, 63-68.

Walker, H. and Lev, J. *Statistical Inference*. New York: Henry Holt, 1953.

Received December 8, 1961

Manifestations of Anxiety on the Rorschach Test

CHARLES NEURINGER
University of North Dakota

Within the last decade there has appeared a stream of investigations which concerned themselves with how anxiety is manifested on the Rorschach test. A perusal of these studies reveals that the results have been so equivocal, inconsistent, and contradictory that they could easily lead the casual observer into assuming that the Rorschach has a poor capacity to reflect any anxious feelings of individuals taking the test. It is the present author's contention that a more detailed analysis of these reports reveals certain highly suggestive trends linking Rorschach data to anxiety when certain critical experimental conditions are taken into account.

It is felt that the Rorschach test is sensitive to anxiety but that the relationship has been obscured by the various experimental maneuvers and operational definitions of anxiety used to evaluate the relationship. The obscuration can be traced to three related sources. (1) None of the experimental designs were equivalent in a replication sense, and therefore reliable and consistent results could not be expected when the experimental conditions differed from one study to another. (2) The use of different subject populations used by the varying investigators has confused the search for general signs of anxiety on the Rorschach. Besides the use of "normal" adults, investigators have used children, adolescents, college undergraduates, neuropsychiatric patients, individuals undergoing brain surgery, fatigued combat flying personnel, etc., as subjects. The manner in which these groups experience, handle, and reflect anxiety on the Rorschach test is quite different, and one should not expect identical or similar anxiety norms from these groups on the Rorschach.

However, the most serious cause of the seemingly contradictory results stems from (3) the varying operational measures of anxiety used by the researchers. Anxiety itself is a diffuse and multidimensional concept and has accrued a variety of meanings. Even though the range of definitions of anxiety was great, it was measured in an even more bewildering number of ways. Electroshock, loud noises, contrived "ego-involving" stress conditions, derogatory comments by the examiner, etc., were used to generate anxiety. Other investigators took advantage of stress provoking conditions (e.g., surgery, distressing physical and psychological examinations, etc.) to investigate the effects of these situations on Rorschach protocols. In other studies, anxiety scales were used to assign subjects to high and low anxiety groups for the purpose of evaluating differences in their Rorschachs. Patients diagnosed as "anxiety reactions" or "anxiety neurotics" were utilized in a similar manner. Anxiety-like states were suggested under hypnosis in one experiment while subjects who had been under long term stress (combat flying) were used in another study.

It seems that the experimental conditions and criterion measures of anxiety used in investigating the relationship between anxiety manifestations and the Rorschach were not identical. It is therefore not surprising that the results of the investigations attempting to link anxiety to the Rorschach test have been inconsistent and contradictory.

However, if one sifts through the maze of differing studies, discarding some which seem to be inappropriate to the problem, some consistent trends appear. In this presentation the varying investigations will be reviewed

and an attempt to relate them to the experimental conditions will be made. The studies can be divided into two broad categories. They are (1) studies of content as determined by the RCT (Rorschach Content Test) Anxiety Scale and (2) investigations of Rorschach determinants.

CONTENT STUDIES WITH THE RCT ANXIETY SCALE

Elizur (1949) using Rorschach content material developed the RCT, a checklist scale that he felt could be used to measure anxiety in a quantifiable manner. A group of neurotic and normal controls were instructed to fill out a self-rating sheet and a questionnaire. The subjects were then given the Rorschach test and interviewed after its administration. The protocols were then analyzed in terms of the RCT Anxiety Scale. The RCT was found to correlate .61 with the questionnaire, .52 with the self-rating sheet and .71 with the interview. In addition, Elizur using the RCT could differentiate the neurotic subjects from the normal subjects at the .01 level of confidence. As an aside, it should be mentioned that it was unfortunate that no inquiry was given as part of the Rorschach administration since such a procedure might have further refined the RCT.

Gorlow, Zimmet and Fine (1952) administered Rorschachs to delinquent and normal adolescents. The RCT was reported to successfully differentiate between the delinquents and the normal adolescents. Westrope (1953) used the MAS to separate out a high and low anxiety group from among college undergraduates. The Wechsler-Bellevue Digit Symbol Subtest with occasional shocks was administered to her subjects. She reported that the RCT discriminated the anxious from the non-anxious subjects successfully. Goodstein (1954) administered the Rorschach and the MAS to college undergraduates and reported that the MAS and the RCT were found to be significantly and

positively correlated at the .01 level. Vernallis (1955) using neuropsychiatric patients who ground or gnashed their teeth (which he felt to be an indication of anxiety) and the RCT found a low but significant correlation coefficient between the two measures. Goodstein and Goldberger (1955) also utilized neuropsychiatric patients in studying the RCT. Using the MAS, they separated out a group of high and low scorers. Their Rorschachs were analyzed by the RCT method, and it was found to discriminate between the high and low anxiety group as measured by the MAS.

On the other hand, Holtzman, Iscoe and Calvin (1954) utilized the MAS to separate out high and low anxiety undergraduate female groups. They report that there was no significant difference in RCT scores between the two groups. However, it should be noted that in this experiment only four Rorschach cards were used. They were cards II, III, VIII and IX, both in their standard forms and as achromatic reproductions. Since only four cards were used, the results of this study should not be considered as an adequate evaluation of the RCT. Kates and Schwartz (1958) also utilized two groups of female undergraduates arranged in high and low anxiety groups, as measured by the MAS. The subjects were given either a typewritten personality evaluation calculated to generate stress or a stress provoking questionnaire. The investigators found that the RCT did not discriminate the high from the low anxiety groups. In fact, they report that the high anxiety group was characterized by lower RCT scores than the low anxiety group. Lucas (1961) administered the Rorschach to nine year old normal children and then eight weeks later put one group through a stress provoking situation. They were frustrated in their ability to do a simple ball and cue problem by experimental means (accompanied by depreciating statements from the examiner). Lucas reported that the ex-

perimental group did not have a significantly higher RCT score than the control group after a second administration. In fact, the RCT scores decreased in both groups on the second administration. The use of children in this experiment may have limited the discriminatory power of the RCT. It would be an unfair test of the RCT to use it on children whose maturational and developmental patterns are quite different from that of the adult's from which the RCT norms were developed.

RORSCHACH DETERMINANTS UNDER VARYING CONDITIONS

The relationship between anxiety and Rorschach determinants has been investigated under several different kinds of conditions. They are (1) laboratory induced anxiety, (2) anxiety scale scores, (3) anxiety scale scores coupled with laboratory induced stress, (4) anxious psychiatric populations, (5) situational stress, and (6) a miscellaneous group. These are broad categories, and it should be remembered that there was variation within each grouping in terms of the kinds of subjects utilized, types of stress conditions provided and design of the experiments.

Laboratory Induced Anxiety.

In this series of studies the general method of investigation used was to divide subjects into two groups. An experimental group usually received some kind of stress treatment while the control group did not. Three studies are reported here. The first one, done by Eichler (1951), was a pivotal one since many other experimenters have tried to relate their findings to his. Eichler matched male undergraduates with the Behn Rorschach cards and on the basis of the results set up a matched experimental and control group. The experimental group subjects were given subtraction problems and received a shock if they produced a wrong answer. Then both groups had the Rorschach adminis-

tered to them. The experimental subjects were strapped into a chair and told that if they gave a "wrong" answer on the Rorschach they would be shocked. The subjects were not told what constituted a wrong answer, and in actual fact they were not shocked at all. Eichler's results indicate that the shocked groups gave more surface shading responses, less W responses, fewer total number of responses, more oligophrenic details, less F responses, rejected more cards, gave fewer popular responses, and fewer responses involving color than did the control subjects.

Allerhand (1954) substantiated Eichler's findings concerning increased surface shading responses under shock conditions. He had his experimental group shocked while they tried to solve a key manipulation problem. Stopol (1954) had undergraduates do the Wechsler-Bellevue Digit Symbol Subtest under (1) task distraction stress and (2) failure stress in which the experimenter constantly disapproved of the subject's performance. He found no significantly distinguishing differentiating characteristics reflected on the Rorschach for the two types of stress. Unfortunately, he did not have a non-stress control group, and the results of his study have applicability only to differences between certain kinds of stress.

Anxiety Scale Scores.

In this series of studies, scales that purport to measure anxiety were used to identify those subjects who were "anxious". In all but one of the studies reviewed here, the Taylor Manifest Anxiety Scale (MAS) was used as the criterion measure of anxiety. One study utilized Cattell's IPAT anxiety scale. The study by Goodstein (1954) was cited previously in the section concerning the RCT. In that study, he also correlated the MAS with the total number of responses on the Rorschach. He reported a low but positively significant relationship between the two measures. The study by

Holtzman, Iscoe and Calvin (1954) was also cited previously. Using four Rorschach cards in a chromatic and corresponding achromatic version, they found little differences in number of responses, F+%, Sum C, A or H content and number of card rejections between their high and low MAS scorers. Again the results of this particular study must be evaluated in the light of their particular experimental conditions (i.e., only four cards were administered in an unusual manner).

Goodstein and Goldberger (1955) used the MAS to separate out a high and low anxiety group of neuropsychiatric patients. Their Rorschach protocols showed that the high MAS scorers gave fewer W responses and fewer color responses than the low MAS scorers. Levitt (1957) administered the Children's MAS and the Rorschach to children in a clinic and to normal children. He reported that the CMAS scores did not correlate with any of the Rorschach determinants studied. He also found greater W%, A%, more Hd responses, M responses, Sum C, Sum Y, CF and YF for the clinic children as compared to the controls. Wagner (1961) used the IPAT as his measure of anxiety and administered it along with the Rorschach to college undergraduates. He reported a positive correlation between IPAT scores and the presence of both aggressive movement responses and anatomy responses only when they appeared together in a protocol.

Anxiety Scale Scores Coupled with Laboratory Induced Stress

This series of studies constituted a refinement over the use of anxiety scales alone. These investigators introduced some sort of experimental stress into the situation which was calculated to activate anxiety reactions. This allowed the researcher to observe the manner in which the low and high anxiety subjects reacted to the stress. Westrope's (1953) study

has already been described in the section dealing with the RCT. She also reported higher surface shading responses for her MAS high scorers under shock condition as compared to the control subjects. Cox and Sarason (1954), after using the MAS to divide undergraduates into a high and low anxious group, administered both a stressful Rorschach (curt instructions, telling the subject that this was a test of intelligence, etc.) and a kindly and relaxed one. They reported that the high MAS scorers gave fewer W responses, fewer populars, more subjective, personalized and self-centered responses, more irrelevant verbalizations, more M, FM and m responses, more F— responses, more personal body responses (sex and anatomical), more diffuse shading responses (FK and K), more form secondary responses (K, C', C'F, c, cF, CF and C), fewer form primary responses (FK, FC', Fc and FC), more perseverations and fewer surface shading responses (Fc, cF and c) than the control subjects. However, it should be noted that c and cF were found to be given more often by the experimental subjects under the form secondary response category than by the control subjects.

Schwartz and Kates (1957) divided undergraduates into a high and low anxiety group by the use of the MAS. They took great pains to find subjects who were free of any kind of environmental stress at the time that the study was conducted. All the subjects were given a Rorschach and then two groups, each comprising high and low MAS scorers, were used as the experimental and control groups. The experimental group subjects were all given the same personality evaluation advising them that they were poorly adjusted. Later on both groups received a second Rorschach examination. Schwartz and Kates concluded that fewer W responses, more m responses and fewer surface shading responses seemed to be a sign of the effects of stress re-

ardless of anxiety level. Increased M responses appeared more often in the low anxiety group as compared to the high anxiety subjects. Longer reaction times, fewer M responses and more F responses characterized the high MAS group as compared to the low anxiety group. Kates and Schwartz (1958) in another study using a similar design (already described in the section on the RCT) found that the number of responses decreased between the first and second administration of the Rorschach for their stress group and that the high anxiety group had fewer M, FM and m responses than the low anxiety group.

Anxious Psychiatric Populations.

In these studies, neuropsychiatric populations specifically diagnosed as either anxiety reactions or anxiety neurotics were compared with other groups. Kates (1950), comparing the Rorschach protocols of patients diagnosed as anxiety reactions with patients diagnosed as obsessive-compulsive reactions, found that the anxiety group gave fewer responses, rejected more cards, gave less rare details, fewer M responses, used more FC, CF and C determinants than the obsessive-compulsive patients. More of the anxiety patients used F exclusively in their records than the obsessive-compulsive patients. The kinds of differences found in this study should be understood in terms of the two groups being compared. It is not surprising that the obsessive-compulsive patients gave more responses and used more rare details than the anxiety reaction subjects. Krasner and Kornreich (1954) compared the Rorschachs of anxiety neurotic patients with normal controls and reported that fewer responses, longer reaction times, more d, and H responses were given by the anxious group. Fewer m and more popular responses also characterized the anxious group as compared to the normal controls. It should not be surprising that the differences ob-

served between the anxiety neurotics as compared to the normal subjects should be greater than the differences found between the anxiety reactions as compared to the obsessive-compulsive population, since many aspects of neurotic functioning are shared by the anxiety reaction group and the obsessive-compulsive subjects and therefore reflected on the Rorschach. How much of the results are due to neurotic functioning itself and how much can be attributed to anxiety is a difficult matter to decide.

Situational Stress

These studies are characterized by the experimenters trying to make use of real life stress situations because it was felt that these conditions were more representative of the kinds of stress that people encounter, and therefore Rorschach protocols should reflect a truer picture of anxiety than the data derived from laboratory experiments.

Alexander and Ax (1951) report a study in which they examined combat airmen who were under great fatigue and stress conditions. They report that fewer responses, poorer form level, verbalizations of threatening content, fewer M responses and fewer color responses characterize this group. Berger (1953) felt that neuropsychiatric patients just coming into a hospital situation would be more anxious than patients that had been hospitalized for over six months. Rorschachs were administered to incoming patients and to patients that had been in the hospital for about six months. He found that the incoming patients gave fewer responses, more F+ responses, had longer reaction times, fewer contents, more W responses, less Sum C, Sum Y, Sum V, greater number of animal responses, more card rejections and fewer populars than the six months control group. The incoming patient group was retested six weeks later, and no differences could be found between the results of the second administra-

tion and the previously used six month control group.

Fisher (1958) conjectured that a gynecological examination given to female neuropsychiatric patients was psychologically and physically very anxiety provoking. An experimental group received Rorschachs after a gynecological examination had been performed on them. A control group received the Rorschach but not the gynecological examination. Five days later both groups received a second administration of the Rorschach. Fisher's raters could find little difference on the Rorschach between the first and second administration for both groups. It is unfortunate that Fisher did not employ a normal control group since the effects of psychosis and the possible presence of a constant low level of anxiety may have obscured real differences in how women react to gynecological examinations. Schon and Bard (1958) studied the Rorschachs of female patients undergoing hypophysectomy (removal of the pituitary gland) or mastectomy (removal of the breast) as surgical means of arresting metastatic breast cancer. The Rorschachs were administered before and after surgery. The investigators report that the hypophysectomy group gave shorter reaction times, fewer F responses and more F- responses after surgery as compared to the mastectomy group. However, the mastectomy group gave fewer W responses. It should be noted here that the results of the hypophysectomy group could well be confounded by the kinds of intellectual changes one could expect from traumatic surgical intervention on brain tissue and may not totally reflect the effects of anxiety about surgery.

Miscellaneous Studies.

Rabinovitch (1954) felt that CF, C, C', KF and c responses were indicators of anxiety and that populars and F+ responses were non-anxiety indicators. He administered Ror-

schachs to neuropsychiatric patients and later had their verbalizations put on to slides and tachistoscopically presented to them. He found that for words used as part of his anxiety indicator signs, GSR deflections and recognition thresholds were higher than for his non-anxiety indicators.

Levitt and Grosz (1960) decided that subjects should be their own controls in anxiety experiments. They attempted to do this by inducing anxiety via hypnotic suggestion. Rorschachs were taken from medical and nursing students in waking states, hypnotic states, and anxiety induced hypnotic states. The anxiety was induced by giving detailed suggestions involving such words as, "anxiety", "fear", "apprehension", and "panic". They found that W, F+, Sum Y, M- and low reaction time were factors that reflect anxiety. The results of this study was predicated on the assumption that anxiety was induced under the hypnotic state. However, it should be pointed out that both authors had some difficulty in agreeing about the degree of anxiety that the subjects showed in the anxiety induced hypnotic state.

EVALUATION

After carefully evaluating the studies reviewed here, several trends emerged concerning the sensitivity of the Rorschach to anxiety. The results are only confusing when they are lumped together regardless of the operations used to measure anxiety. In fact, several of the studies cited here (Fisher, 1958; Holtzman, Iscoe & Calvin, 1954; Schon & Bard, 1958; Stopol, 1954) have either methodological weaknesses or were designed primarily to study some other aspect of the Rorschach and have only secondarily touched on the problem of anxiety. The results of these studies tended to hinder efforts to clarify the relationship between anxiety and the Rorschach test. Accordingly, these studies did not play very much of a role in the analysis of the experimen-

tal conditions linking anxiety manifestations to Rorschach data.

The important key to how anxiety is reflected on the test involved a linking of Rorschach results to the kinds of anxiety generated by the different investigators. It would seem obvious that the anxiety experiences of combat flyers are not the same as those felt by students who, while receiving mild electric shocks when doing subtraction problems, knew that the pain would cease as soon as they leave the testing room. The conclusions reached here will be presented in two parts. They are (1) results of the study of content through the RCT Anxiety Scale and (2) findings from the analysis of Rorschach structural determinants.

The RCT Anxiety Scale.

The RCT is essentially a device that utilizes the content material of the Rorschach, and it appears to be a trustworthy instrument under a number of different conditions. It seems to be less effective when used with children. There are aspects of children's general reactions to the Rorschach that would have predicted this.¹ Content material seems to be one aspect of the Rorschach that appears to be sensitive to the different kinds of anxiety experienced by undergraduates, neuropsychiatric patients and adolescents. The communication of anxiety affects through content appears to be a basic and immediate process and seems to be easily available to most people. Although content responses were not studied in as much detail as the structural determinants, these studies did report the presence of anatomy, sex, animal and threatening contents among the anxious subjects. Unfortunately for scientific purposes, judgments about content are more subjective and prone to individual differences in interpretation than the allegedly more

precise structural determinants. A great deal has been written about content in an intuitive and clinical manner (Beck, 1952; Phillips and Smith, 1953; Schafer, 1954; Piotrowski, 1957) but there has been little programmatic research done. Content may be the best clue for assessing anxiety, but the difficulties in deciding what constitutes an anxiety indicating content are many. However, the RCT does seem to offer some help.

Structural Determinants.

As far as Rorschach determinants go, they do reflect anxiety, but they do it differently under different conditions. Utilizing those studies which seem to adequately deal with the problem, the following trends evolved.

(a) It was found that two clusters of linkage between experimental conditions and results emerged and (b) that the reactions to anxiety by children are different than those elicited from adults.

(a) When highly intelligent and verbal undergraduates are used as subjects, either under laboratory induced stress alone or coupled with an assessment by a scale purporting to measure anxiety, the trends seem to point in a direction quite different from the one found for people tested under a real life, long term situational stress that is impinging on their lives. The results of the former group of investigations (laboratory induced stress and/or high scores on anxiety scales for college undergraduates) points to a decreased number of responses, increased surface shading responses, fewer W responses, fewer populars, and more M and m responses. It would appear as if that kind of anxiety has mobilized the subjects, and they are forced to fall back and vigilantly assess the blot material. They seem to strive for tighter control of their responses at the price of more global organization (less R, less W, fewer P responses), and they become very careful and cautious in

¹A fuller discussion of this point is to be found in the section dealing with structural determinants.

scrutinizing the blot material (more surface shading, more M and m). It is as if they have become very sensitive to nuances in the blot material.

The subjects caught in real life anxiety provoking situations react much differently. They also give fewer responses, but they reject more cards, give more F responses, produce less M, m and color responses. They have longer reaction times and give more popular responses. It appears that they are trying to avoid coming to grips with the blot material (fewer responses and more rejections). They give mundane responses that do not demand very much involvement on their part (more F, more populars, less M, m and color responses). If the other subjects appear to have become oversensitized, these subjects appear to be insensitive.

These subjects seem to be denying themselves the opportunity to deal with their perceptions and feelings and prefer to escape into denial and avoidance. They seem to have been drained and immobilized by their kind of anxiety and have become more undifferentiated in their psychological organization.

(b) Constriction of some sort seems to be a general reaction to anxiety. Both groups tend to give fewer responses but for wholly different reasons. Of the different groups studied here, only children did not show coarctation in the presence of anxiety. Anxiety seems to have a liberating effect, in the sense that children tend to produce many kinds of responses while under stress. This occurred both in the Levitt (1957) and Lucus (1961) investigations and could have been predicted from our knowledge of child development. This trend has already been observed by Ames *et al.* (1952) and Klopfer *et al.* (1954). Children have not yet developed settled patterns of reacting, and their behavior may go in any number of directions under stress.

The reader must remember that the preceding conclusions about

Rorschach clues to anxiety represent distillates from the studies reviewed here. Not all the studies conform exactly to the findings expressed here. There are too many highly specific variations in the designs to be able to make more embracing statements. The conclusions reached here should be considered as only highly suggestive trends, but they do indicate that the Rorschach can reflect anxious feelings. The important thing to remember is that Rorschach results must be linked to anxiety with an appropriate operational subscript in order that the results be meaningful to others. The experimenter in the laboratory doing basic research on Rorschach determinants cannot expect to find the kinds of responses that the examiner encounters when dealing with tremulous patients in a clinic. Likewise, the clinician should not be disappointed if he does not find the anxiety indicators that are produced in the laboratory by highly articulate and intelligent undergraduates.

SUMMARY

Many studies have appeared which have attempted to relate the Rorschach test to anxiety. On the surface, the results of these studies appeared to be equivocal. Three causes of confusion were noted. They were (1) non-replication of experiments, (2) the use of differing subject populations and (3) the presence of varying operational measures of anxiety. It was the author's contention that the Rorschach test could be shown to be sensitive to anxiety when the results were linked to certain kinds of experimental conditions.

The literature linking the RCT Anxiety Scale and Rorschach determinants under a varying set of experimental designs was reviewed, and it was concluded that (1) the RCT was a helpful method of assessing anxiety, by the use of contents, under a number of experimental conditions. It was also suggested that (2) the complex of Rorschach determinants linked to

laboratory induced stress, with or without anxiety scale score measures as anxiety criteria for college undergraduates, was quite different from the cluster of determinants found for individuals suffering from long term or situational stresses.

REFERENCES

- Alexander, L. & Ax, A. F. Rorschach studies in combat flying personnel in Hoch, P. H. & Zubin, J. (Ed's). *Relation of psychological tests to psychiatry*. New York: Grune & Stratton, 1951.
- Allerhand, M. E. Chiaroscuro determinants of Rorschach test as indicators of manifest anxiety. *J. proj. Tech.*, 1954, 18, 407-413.
- Ames, Louise B. et al. *Child Rorschach Responses*. New York: P. B. Hoeber, 1952.
- Beck, S. J. *Rorschach's Test. Vol. III-Advances in interpretation*. New York: Grune & Stratton, 1952.
- Berger, D. The Rorschach as a measure of real life stress. *J. consult. Psychol.*, 1953, 17, 355-358.
- Cox, F. N. & Sarason, S. B. Test anxiety and Rorschach performance. *J. abnorm. soc. Psychol.*, 1954, 49, 371-377.
- Eichler, R. M. Experimental stress and alleged Rorschach indices of anxiety. *J. abnorm. soc. Psychol.*, 1951, 46, 344-355.
- Elizur, A. Content analysis of the Rorschach with regard to anxiety and hostility. *J. proj. Tech.*, 1949, 13, 247-284.
- Fisher, Rhoda L. The effects of a disturbing situation upon the stability of various projective tests. *Psychol. Monogr.*, 1958, 72, 1-23.
- Goodstein, L. D. Interrelationships among several measures of anxiety and hostility. *J. consult. Psychol.*, 1954, 18, 35-39.
- Goodstein, L. D. & Goldberger, L. Manifest anxiety and Rorschach performance in a chronic patient population. *J. consult. Psychol.*, 1955, 19, 339-344.
- Gorlow, L., Zimmet, C. N. & Fine, H. J. The validity of anxiety and hostility Rorschach content scores among adolescents. *J. consult. Psychol.*, 1952, 16, 73-75.
- Holtzman, W. H., Iscoe, I. & Calvin, A. D. Rorschach color responses and manifest anxiety in college women. *J. consult. Psychol.*, 1954, 18, 317-324.
- Kates, S. L. Objective Rorschach response patterns differentiating anxiety reactions from obsessive-compulsive reactions. *J. consult. Psychol.*, 1950, 14, 226-229.
- Kates, S. L. & Schwartz, F. Stress, anxiety and response complexity on the Rorschach test. *J. proj. Tech.*, 1958, 22, 64-69.
- Klopfer, B. et al. *Developments in Rorschach techniques. Vol. II*. New York: World Book Comp., 1954.
- Krasner, L. & Korneich, M. Psychosomatic illness and projective tests: The Rorschach test. *J. proj. Tech.*, 1954, 18, 355-367.
- Levitt, E. E. Alleged Rorschach anxiety indices in children. *J. proj. Tech.*, 1957, 21, 261-264.
- Levitt, E. E. & Grosz, H. J. A comparison of quantifiable Rorschach anxiety indicators in hypnotically induced anxiety and normal states. *J. consult. Psychol.*, 1960, 24, 31-34.
- Lucas, Winifred B. The effects of frustration on the Rorschach responses of nine year old children. *J. proj. Tech.*, 1961, 25, 199-204.
- Phillips, L. & Smith, J. G. *Rorschach interpretation: Advanced technique*. New York: Grune & Stratton, 1953.
- Piotrowski, Z. A. *Perceptanalysis*. New York: McMillan, 1957.
- Rabinovitch, M. S. Physiologic response, perceptual threshold and Rorschach test anxiety indices. *J. proj. Tech.*, 1954, 18, 379-386.
- Schafer, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- Schon, Martha & Bard, M. The effects of hypophysectomy on personality in women with metastatic breast cancer as revealed by the Rorschach test. *J. proj. Tech.*, 1958, 22, 440-445.
- Schwartz, F. & Kates, S. L. Rorschach performance, anxiety level and stress. *J. proj. Tech.*, 1957, 21, 154-160.
- Stopol, M. S. Rorschach performance in relation to two types of stress. *J. consult. Psychol.*, 1954, 18, 11-15.
- Vernallis, F. F. Teeth-grinding: Some relationships to anxiety, hostility and hyperactivity. *J. clin. Psychol.*, 1955, 11, 389-391.
- Wagner, E. E. The integration of aggressive movement responses and anatomy responses on the Rorschach in producing anxiety. *J. proj. Tech.*, 1961, 25, 212-215.
- Westrope, Martha R. Relations among Rorschach indices, manifest anxiety and performance under stress. *J. abnorm. soc. Psychol.*, 1953, 48, 515-524.

Received March 1, 1962

Problems in the Measurement of Aggression-Anxiety

PAUL ROTHBAUS
VA Hospital, Houston, Texas

The catharsis hypothesis (Dollard, Doob, Miller, Mowrer and Sears, 1939, p. 50) asserts that the making of an aggressive response reduces the instigation to aggression. The validity of this hypothesis remains in doubt despite widespread investigation. As Berkowitz (1958) points out in his review of the catharsis literature, one question consistently clouds interpretation of results. Specifically, does aggression-anxiety ensue after a cathartic expression of hostility, masking and suppressing further hostile expression, and so creating an illusion that the underlying hostility has been reduced? Such a possibility makes accurate measurement of aggression-anxiety crucial for research concerning the reduction of hostility. This report comments on certain problems involved in one such measurement attempt.

McClelland (1951) proposed a measure of aggression-anxiety in his discussion of a clinical case. He suggested that high aggression-anxiety causes the aggression in patients' Thematic Apperception Test protocols to originate from impersonal sources, e. g., people die for unknown reasons, or suffer injury from impersonal agents such as windstorms and disease. McClelland had in mind inhibitory fear or aggression-anxiety of a characterological nature, stemming from severe punishment for childhood aggressions.

The Thematic Apperception Test (TAT) has also been used as a measure of transitory states of aggression-anxiety. Stone (1950) compared football players and control Ss both during and after the football season. No differences were found on the TAT during the football season, but afterwards the football players showed sig-

nificantly less manifest aggression. The aggression which did appear in their protocols was of an indirect or impersonal nature. Stone interpreted this finding as indicating that the football players felt guilty (aggression-anxiety) about their athletic aggression, and accordingly were inhibiting direct expression of aggression in the TAT.

The discussion in this paper is directed toward the measurement of aggression-anxiety using the TAT in experimental research on hostility reduction, where the aggression-anxiety is probably transient, as in the Stone study, rather than characterological, as in the McClelland case report. Two problems will be examined; (a) the effects on judges due to prolonged scoring of numerous records, and (b) the unit of analysis for scoring, important because the typical method of TAT administration allows S freedom to give stories of any length. Scale scores may be confounded with test length (similar difficulties have appeared with the Rorschach). Perhaps a better method than scoring all instances of aggression is obtained when a TAT card is scored for either the presence or absence of aggression. Both methods of scoring are compared in this paper.

METHOD

Six TAT cards (Murray, 1943) were administered by slide projector to 236 college men and women in 12 groups of about 20 Ss each. Ss were provided with one sheet of paper for each card, and asked to write a story telling what was happening, how it began and how it would turn out. Slides of cards 3 BM, 6 GF, 7 BM, 8 BM, 12 BG, and 14 were used.

An initial 33 TAT protocols were scored independently and in the same order by the author and a graduate

student assistant employing Clark's criteria (1955). The first six of these 33 protocols were then rescored. Subsequently, all 236 records were scored for manifest and impersonal aggression employing a modification of Clark's criteria. While using the modified criteria, two methods of scoring were studied; (a) rating a TAT story for the sum of all instances of aggression, (the score is then the total instances of aggression in all six TAT stories), (b) rating a TAT story for the presence or absence of aggression, (the score is the total number of stories in which aggression occurred).

Modified Criteria for Manifest Aggression. 1. One individual intentionally inflicts physical harm or injury upon another. 2. One individual scolds, offends, insults, belittles, or over-bearingly dominates another. 3. Individuals quarrel, or have disagreements characterized by anger, hatred, resentment, tension or oppression. 4. One individual intentionally causes personal tragedy, misfortune, unhappiness or discomfort to another; e.g., (a) Her parents tell her that under no condition may she attend the dance. The girl is heartbroken. (b) Because of his superior's report, he loses his job. (c) His father's insistence that he learn a trade makes him unhappy, for he has no interest in it. (d) His parents refuse to lend him the family car. The girl finds another date, and he loses her. 5. Individuals commit crimes, or engage in destructive acts, or lose their tempers and perform some violent act. Punishment by law for crimes committed, however, is not scored as manifest aggression. 6. An individual commits suicide.

Modified Criteria for Scoring Impersonal Aggression. 1. The agent of the aggression is impersonal, e.g., (a) Injury or harm by animals, germs, creatures from outer space, or spirits, (b) Injury due to "acts of God", such as storms, floods, landslides, lightning and earthquakes, (c) Injury due to accidents such as drowning, ship-

wrecks or falls, (d) Injury or death due to natural causes, such as sickness, old age, or childbirth mishaps, (e) Personal misfortunes and/or tragedies not caused by another individual. 2. The agent of aggression is unspecified, or the cause of harm is unspecified, vague, or unknown, e.g., "Years later the mother returns, but her daughter has already died," or "In the darkness, something attacks him from behind, and he doesn't awaken until days later." 3. Although a person is the agent of aggression, intent to harm is absent, e.g., "His brakes fail and the car kills the boy," or "He does not know the gun is loaded and it goes off and hits his father."

RESULTS

Differences between judges. Clark obtained interjudge reliabilities of .93 for manifest aggression and .88 for impersonal aggression. In this study the scoring of 33 six-card protocols for all instances of aggression (Clark's scoring method) yielded product moment reliabilities of .94 on manifest aggression and .83 on impersonal aggression.

One notes that the interjudge reliabilities are slightly reduced under the modified criteria. This may be a consequence of the elimination of a spuriously inflating order effect, rather than a true reduction of interjudge

TABLE I—Interjudge Reliabilities
Under Different Systems of
Scoring Aggression

	N	Aggression	
		Manifest	Indirect
Clark Study			
Scoring all Instances	77	.93	.88
Current Study			
Using Clark's Criteria;			
Scoring all Instances	33	.94	.83
Current Study			
1. Using Revised			
Criteria;			
Scoring all			
Instances	236	.88	.78
2. Using Revised			
Criteria;			
Scoring for presence			
or Absence of			
Aggression	236	.86	.73

TABLE II—Increases in Total TAT Aggression Reported by two Judges from the First to the Second Scoring of Six Protocols

TAT Protocol	Judge A Total Aggression			Judge B Total Aggression		
	First Scoring	Second Scoring	Increase	First Scoring	Second Scoring	Increase
S1	8	10	2	6	8	2
S2	7	8	1	6	7	1
S3	18	19	1	13	16	3
S4	6	6	0	5	6	1
S5	9	10	1	4	9	5
S6	3	3	0	1	3	2
Mean Aggression	8.50	9.33	.83	5.83	8.17	2.33

agreement, a point which will be examined later.

Upon rescoring the first six of the 33 TAT protocols, both judges reported significantly more total aggression (Table II). For Judge A, the mean increase per subject in the number of aggressive responses scored was .83 ($t = 2.71, p < .05$). For Judge B, the mean increase was 2.33 ($t = 3.53, p < .02$).

Table II also shows that the judges' increases in the scoring of aggression were unequal. The mean difference in their increase was 1.50 ($t = 2.42, p < .07$). Again referring to Table II, one sees that Judge B (who is a woman) initially scored less aggression than A (mean difference equals 2.67, $t = 3.51, p < .02$) but increased with time more than A did. The mean difference in their ratings of aggression at the time of the second scoring is smaller (1.17) but still significant ($t = 2.44, p < .07$).

Comparison of the two systems of scoring. Table III shows the relationship between the two methods of scoring for both judges, and for the final scores decided upon by the judges under each method.

One notes that here, as in Table I, the correlations for impersonal aggression are lower than those for manifest aggression. The former scale is apparently less reliable and subject to greater error than the latter.

DISCUSSION

Effects of repeated scoring. Whatever the explanation behind the scor-

TABLE III—Correlation Between Two Systems of Scoring Aggression on TAT Cards

(Scoring All Instances of Aggression, and Scoring for the Presence or Absence of Aggression)

	Aggression	
	Manifest Scale	Impersonal Scale
Judge A	.85	.79
Judge B	.87	.78
Consensus Judgment*	.85	.77

*The final score agreed upon by both judges after discussion of the differences in scoring a protocol.

ing differences of the judges (sex differences, personality differences, etc.) two cautions should be observed in view of the fact that judges may report increasing amounts of aggression with repeated scoring.

First, when an experimental design is comparing two methods for reducing hostility, the TAT protocols arising from the two methods must not be scored in blocks, one after the other. The method scored second would have the higher thematic aggression score merely as a consequence of the order of scoring. When comparing two methods an ABBA order of scoring could be followed with the assumption that the increases of the judges were a linear function of repeated scoring. When more experimental treatments or methods are being compared, counterbalancing the order effect will be more complicated. However the ideal situation would exist if judges showed no such increases at all. When judges can vary in scoring identical protocols solely

because the protocols are scored at different times, error variance is introduced which reduces the power of the experimental design to reveal differences between the methods. Criteria which provide many examples should help prevent the order effect. Even so, judges should be given a prolonged period of scoring practice protocols before rating records from the experimental design. Doing so increases the likelihood that the judges will have reached an asymptote by the time the experimental protocols are rated.

The second caution pertains to the interpretation of interjudge reliability. If two judges score protocols in the same order and their scores are increasing due to the order effect, their interjudge reliability correlation will be inflated. Once the order effect "levels off" or reaches an asymptote, the interjudge reliability will not be magnified by the error variance arising from repeated scoring. Calculation of the interjudge reliability at this later time will probably yield a reasonably conservative coefficient.

Comments on the two methods of scoring. In evaluating the two methods for scoring aggression (a) Scoring all instances of aggression, and (b) Scoring only for the presence or absence, the following summary points seem appropriate:

1. The second method offers slightly lower interjudge reliabilities.
2. The two methods are highly correlated, so that empirically the latter system is probably heir to the same difficulties as the original, that is, subject to the same confounding with response length.
3. The latter system does not in fact control the number of responses an S may give, which is the crux of the problem. It merely limits the total score an S may receive. The longer the story an S creates, the greater the likelihood that the theme will contain aggressive content, hence the confounding problem persists.
4. Finally, scoring each card for either the presence or absence of ag-

gression probably constitutes a reduction in the sensitivity and validity of the test.

The problem of the confounding of test length (the total number of responses) with scale scores has plagued projective tests. The Rorschach in particular has been vulnerable in this respect. Holtzman (1961, p. 8) has attempted to control the confounding effect of a variable number of responses to blot material by explicitly defining the number of responses the S may give to a card. It is less clear how such control can be maintained with TAT material, especially under conditions of group testing using a slide projector. Accordingly, when expediency necessitates group TAT testing, research data should be examined to determine whether aggression-anxiety scores were correlated and confounded with test length. Where such correlations exist the following question must be weighed: Have differences in experimental treatments caused or motivated Ss under different conditions to write stories of varying lengths, the scoring of which gives rise to aggression scale differences? The increased productivity might be a fact of more pervasive importance than any of the incidental scale differences.

SUMMARY

TAT stories of 236 college men and women were scored by two judges for manifest and impersonal aggression. There was a significant tendency for both judges to increase, after repeated scoring, in the amount of aggression attributed to protocols. Such increases necessitate counterbalancing or randomizing the order of scoring by all judges of protocols in experimental designs comparing various treatments. Increases can also create spuriously high interjudge reliabilities if protocols are scored in the same order by judges. Criteria with explicit examples of scorable responses probably help reduce the increasing order effect.

Consideration was also given to two methods for scoring: (1) scoring each instance of aggression, and (2) scoring a story for the presence or absence of aggression. The first method offered superior interjudge agreement. Both methods yield scale scores which may be subject to confounding with test length. Experimenters are cautioned to explore whether scale scores are related to test length before making interpretations of statistical results.

REFERENCES

- Berkowitz, L. The expression and reduction of hostility. *Psychol. Bull.*, 1958, 55, 257-283.
Clark, R. A. The effect of sexual motivation on fantasy. In D. McClelland (ED.), *Studies in Motivation*. New York: Appleton-Century-Crofts, 1955.
Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. *Frustration and aggression*. New Haven, Yale Univ. Press, 1939.
Holtzman, W. H., Thorpe, J. S., Swartz, J. D., Herron, E. W. *Inkblot Perception and Personality*. University of Texas Press, Austin, 1961.
McClelland, D. C. *Personality*. New York: Dryden, 1951.
Murray, Henry A. *Thematic apperception test and manual*. Cambridge, Mass.: Harvard Univer. Press, 1943.
Stone, A. A. The effect of sanctioned overt aggression on total instigation responses. Unpublished honors thesis, Harvard Univ., 1950.

Received January 9, 1962

The Effect of Preliminary Verbal Conditioning on Inkblot Test Responses¹

HARRY SACKS²

University of California at Los Angeles

There has been much discussion and considerable research on the topic of how stable an instrument an inkblot test is in its task of assessing personality. However, most of this research deals with such topics as the effect of different examiners on a subject's inkblot test responses (Baughman, 1951; Gibby, 1952; and Lord, 1950), the effect of unconscious cues by the examiner (Wickes, 1956), the effect of varying the instructions (Fosberg, 1941; Fosberg, 1938; Gibby, 1951; Hutt, 1949; and Hutt, 1950) and other similar research in which the test is not presented in the same standardized fashion for all subjects. Such studies are interesting and serve to emphasize the fact that an inkblot test, like other psychological tests, must be given under standardized conditions with standardized instructions if test protocols of different subjects are to accurately reflect differences between subjects rather than artifacts and "errors" of the testing situation.

This study explored the extent to which happenings of the day on which the test was taken could affect a subject's performance on an inkblot test if the test were administered in a carefully standardized fashion; specifically, to what extent such happenings would affect the number of human movement responses (M) given to the Wickes Inkblots (Wickes, 1956). The Wickes Inkblots were

chosen because the percentage of human movement responses which each card elicits is known, there are two sets of inkblots which are matched for the percentage of human movement responses which they elicit, and no card elicits less than twenty per cent human movement responses as determined by Wickes' original study. For example, if a subject read material involving human movement before he took the inkblot test, would this increase the number of human movement responses which were given in the test? Or, to take an extreme situation, if the subject were actually conditioned to show preferences for human movement responses on another task just before taking the inkblot test, would this result in the subject's producing more human movement responses on the inkblot test itself, as it is fairly well established that there is stimulus generalization from words to the objects which they represent? Therefore, if subjects would be conditioned to choose human movement phrases, would this carry over to a new task causing them to perceive more human movement in ambiguous inkblots, in view of the fact that phrases describing human movement have been significantly raised in the subjects' habit hierarchy and the subject would have the opportunity to continue to choose similar responses on the inkblot test? If happenings of the day of the testing could significantly affect such a crucial and presumably stable variable as the number of human movement responses which a subject produces, then the stimuli with which the subject comes into contact just before taking the test should be controlled or at least be within the awareness of the examiner

¹Based on a study submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of California at Los Angeles. The writer wishes to express his appreciation to the members of his committee, Bruno Klopfer, Richard P. Barthol, Allen B. Rosenstein, B. Lamar Johnson, and George F. J. Lechner (chairman).

²Now at the State of California Mental Hygiene Clinic, Covina, California.

in order to increase the accuracy of his test interpretations.

PROCEDURE

The subjects were eighty male volunteers who were patients at Wadsworth V.A. Hospital, a general medical hospital. None of the subjects were diagnosed by their physician as psychotic or brain damaged. Subjects were randomly divided into three groups. Subjects in Group 1 were first given the Shipley-Hartford intelligence test. They were then given the first of Wickes' two sets of inkblots with written instructions which told the subject to write down for each of the fifteen blots the one response which makes the "strongest impression" on him. The examiner was not in the room while the test was taken. When a subject finished with this test, he was taken to another room where he was seated in a booth. This booth had a glass screen on which were projected eighty pairs of phrases, one pair at a time. Each pair consisted of one human movement phrase such as "running boy" and one non-human movement phrase such as "setting sun." The two phrases in each pair had the same number of words. There were also approximately the same number of words ending in "ing" in each group. The position of the phrases on the screen was randomized to prevent a subject from becoming conditioned to the location of the phrase instead of its content. The examiner sat just outside the booth. The subject was instructed to state which phrase in each pair made the "strongest impression" on him. The examiner wrote down whether a human movement or non-human phrase was chosen. While the first twenty pairs of phrases were shown, the examiner remained quiet and merely wrote down the subject's choice. Beginning with the twenty-first pair, however, the examiner attempted to condition the subjects to choose a human movement phrase. The first time the subject chose a hu-

man movement phrase the examiner said, "Good," the second time, "Fine," the third time, "All right." This cycle then began again and continued throughout the remainder of the list. If a subject chose significantly more human movement phrases on the final twenty pairs than on the first twenty, he was considered to be conditioned. The test of significance was the critical ratio between proportions, .10 level, one tailed test. It was necessary to run 42 subjects in order that there would be 20 conditioned and 20 non-conditioned subjects in this group. The other 2 non-conditioned subjects were dropped from the study in order that there be an equal number of subjects under all conditions. When subjects in Group 1 finished viewing the phrases, they immediately took the second set of Wickes' inkblots. The two sets of blots are matched for the number of human movement responses each set elicits, but as an extra precaution, half of each group took one set first and half took the other set first. Group 2 had twenty subjects which went through exactly the same procedure as Group 1 except that the examiner remained silent while the list of phrases were shown. Group 3 went through the same procedure except that instead of viewing the list of phrases, they took the Bender Gestalt between taking the two sets of inkblots. This took approximately the same time as the list of phrases, but did not involve human movement.

The issue was whether or not the conditioned subjects in Group 1 would show an increase in human movement responses on the second inkblot test which would be significantly greater than any increase shown by the subjects in Groups 2 and 3 or the non-conditioned subjects in Group 1.

RESULTS

Table I shows the mean increase in human movement responses for each

TABLE I—Mean Increase in Human Movement Response for Each Condition

Condition	Group	N	Mean Increase
1	1, Successfully conditioned	20	0.25
2	1, Not conditioned	20	0.90
3	2	20	0.65
4	3	20	0.35

of the four conditions in this study. It will be observed that the mean increase in human movement responses for each condition was between 0 and 1, with the *smallest* increase being in those subjects who were successfully conditioned to choose human movement phrases as making the strongest impression on them, while the *largest* increase was shown in those who were not successfully conditioned to the list.

Table II summarizes the analysis of variance of the data. The difference between conditions was decisively non-significant. With an N of 80 subjects the F ratio was only .26. If the F ratio had been ten times as great as that observed in this study it would still fail to be significant at the .05 level.

TABLE II—Analysis of Variance of Increase in Human Movement Responses

Source	df	Sum of squares	Variance Estimate	F
Between Groups	3	5	1.67	
Within Groups	76	481	6.33	.26
Total	79			

Six subjects in this study were completely conditioned, i.e., they chose the human movement response on all of the final twenty choices, which represented an increase of human movement choices from the first twenty pairs that was significant well beyond the .01 level in each case. As a group these subjects gave a total of seventeen human movement responses on the first inkblot test and sixteen human movement responses on

the second inkblot test. Therefore, the subjects who were most strongly conditioned to give human movement responses on the list actually showed a very slight, although obviously insignificant, *decrease* in the number of human movement responses which they produced on the second inkblot test.

It should also be mentioned that in spite of the fact that the subjects in this study included a very wide range of intelligence, intelligence was not a relevant variable in this study. Intelligence did not significantly correlate with conditionability to the list of phrases nor to the number of human movement responses which a subject gave to the initial inkblot test. Even more important to this particular study was the fact that intelligence was not related to increasing the number of human movement responses under any of the conditions in this study.

DISCUSSION

The results of this study strongly support the observations and experimental results of both Rorschach (Rorschach, 1951) and Fosberg (Fosberg, 1941; 1938) who had concluded that the human movement variable on the Rorschach inkblot test was a very stable variable that was not easily susceptible to manipulation. Other studies had not supported this notion (Baughman, 1951; Gibby, 1952; Hutt, 1949; Hutt, 1950; Lord, 1950; Milton, 1949; and Wickes, 1956). However, most of these studies had given the inkblot test under conditions that were not the same for all subjects. Subjects in different groups had been given different examiners, different instructions, different amounts of verbal or non-verbal reinforcements, etc. The majority of studies had, therefore, supported the notion that the human movement variable was not as stable as Rorschach had believed. However, the inkblot tests used in these studies did not represent a projective test in the

true sense of the word. A projective test implies that the test consists of ambiguous stimuli which the subject must then structure according to his own inner resources. However, when the subject is given specific instructions to produce human movement responses, the task is no longer unstructured. The subject has a definite task to perform, namely to produce human movement responses, and to the extent that he is able to follow such instructions the test does not reflect the subject's personality, but the examiner's instructions. The same holds true for the studies in which the examiners reinforced human movement responses while the subject took the test. To the extent that the subject was able, on some level, to perceive the nature of the examiner's reinforcements, the task was structured in terms of following the examiner's implicit instructions to produce more human movement responses rather than for the subject to structure the task in his own manner. It is probable that different examiners get different results to the extent that they give the subject certain cues to respond in a particular manner, even though the examiner, himself, is unaware of emitting such cues. This was the major conclusion of Wickes' study. This present study indicates that the human movement variable on an inkblot test is not readily subject to manipulation by preliminary conditioning of subjects to human movement phrases *provided the test is administered in a standardized fashion*. Conditioning subjects to state that human movement phrases make a "stronger impression" on them did not immediately thereafter cause the subjects to see human movement responses as making a "stronger impression" on them in an inkblot test. It is, therefore, unlikely that human movement responses on an inkblot test will be affected by coming into contact with descriptions of human movement on the day of the testing, in view of the

fact that even conditioning subjects to human movement phrases had no effect on this variable, although the inkblot test immediately followed such conditioning.

A brief explanation should be given about the lack of relationship between intelligence and the number of human movement responses given on the Wickes' inkblots. Unlike the Rorschach, M responses are popular on all of the Wickes blots. It, therefore, requires much less intelligence to see such responses on the Wickes blots than on the Rorschach.

SUMMARY

This study explored the possibility that happenings of the day of the testing might affect inkblot test responses, specifically the number of human movement responses given by a subject on the Wickes inkblots. Eighty subjects were randomly divided into three groups. The first group took an intelligence test, followed by the first set of Wickes' inkblots. The examiner then attempted to condition the subjects to choose human movement phrases as making a "stronger impression" on them than other phrases not involving human movement. There were twenty conditioned and twenty non-conditioned subjects in this group. They then took the second set of Wickes' inkblots to see if there would be a significant increase in human movement responses from the first inkblot test. Two control groups were used. Conditioning had no effect.

REFERENCES

- Baughman, E. E. Rorschach scores as a function of examiner difference. *J. proj. Tech.*, 1951, 15, 243-249.
- Fosberg, I. A. An experimental study of the reliability of the Rorschach psychodiagnostic technique. *Rorschach Res. Exch.*, 1941, 5, 24-36.
- Fosberg, I. A. Rorschach reactions under varied instructions. *Rorschach Res. Exch.*, 1938, 3, 12-30.
- Gibby, R. G. Examiner influence on the Rorschach inquiry. *J. consult. Psychol.*, 1952, 16, 449-455.

- Gibby, R. G. The stability of certain Rorschach variables under conditions of experimentally induced sets: I, The intellectual variables. *J. proj. Tech.*, 1951, 15, 3-25.
- Hutt, M. L., Gibby, R., Milton, E. O., & Pott-harst, K. The effect of varied experimental "sets" upon Rorschach performance. *Rorschach Res. Exch.*, 1949, 13, 240-241.
- Hutt, M. L., Gibby, B., Milton, E. O. & Pott-harst, K. The effect of varied experimental "sets" upon Rorschach test performance. *J. proj. Tech.*, 1950, 14, 181-187.
- Lord, E. Experimentally induced variations in Rorschach performance. *Psychol. Monogr.*, 1950, 64, No. 10 (Whole No. 316).
- Milton, O. The influence of varied experimental sets upon certain Rorschach variables: II, Stability of the human movement variable. Unpublished doctoral dissertation, Univ. of Michigan, 1949.
- Rorschach, H. *Psychodiagnostics*. Grune & Stratton, 1951.
- Wickes, T. A. Examiner influence in a testing situation. *J. consult. Psychol.*, 1956, 20, 23-26.

Received January 4, 1962

Verbal Reinforcement of a TAT Theme¹

WILLIAM L. SIMMONS AND EDWARD G. CHRISTY²
University of Nevada

In recent years many studies dealing with the effect of operant conditioning on response frequency have been reported in the literature. The majority of these studies have employed a positive verbal reinforcer such as "good" (Wilson & Verplanck, 1956) or "mmm-hmm" (Greenspoon, 1955) in order to increase the frequency of varying classes of responses. For the most part, the response units which were reinforced were of relatively short lengths, frequently one word responses, such as plural nouns (Mandler & Kaplan, 1956), adverbs (Wilson & Verplanck, 1956), pronouns (Taffel, 1955), and digits (Tatz, 1956). This line of research has also been introduced into the field of projective testing and incorporated into studies investigating the effect of examiner influence on responses to various projective devices. Of the studies in this area, most have utilized the Rorschach test and have also reinforced short response units, such as responses of movement (Wickes, 1956) and human responses (Gross, 1959). The general conclusion to be drawn from these studies is that an examiner can influence the verbal output of a group of subjects (Krasner, 1958). This would seem to be of special significance in regard to projective testing where the frequency of a given response unit may be seen as having diagnostic value. This is clearly true of the TAT. Many writers (Murray, 1943; Rotter, 1946; Bellak, 1959) have drawn attention to the importance of such an indicant in the understanding of personality dynamics. Thus, an examiner by inadvertently reinforcing a class of responses could increase

their frequency and subsequently make unwarranted inferences on this basis. To date no studies have dealt with the influence of verbal reinforcement on the output of a given TAT theme, and few studies have dealt with response units of greater length than one word or a short phrase as would be required in TAT stories. The present study proposes to investigate this problem and test the specific hypothesis that the frequency of a designated theme on the TAT can be significantly increased as a function of the verbal reinforcements given by an examiner.

METHODOLOGY

Subjects: Forty volunteer students of both sexes, enrolled in classes in personal adjustment at the University of Nevada, were used as subjects.

Procedure: Prior to undertaking the present study, a pilot study was conducted to determine the theme to be reinforced, the TAT cards to be selected, and their order of presentation. On the basis of data obtained from approximately 30 subjects, the theme chosen to be reinforced was that of parent-child interaction. The cards selected and the order of their presentation were as follows: 7GF, 6BM, 7BM, 2, 5, 18GF, 9GF, 3BM, 8GF, and 16. Since it was necessary that a response occur in order for it to be reinforced, it was necessary to select cards that had a high probability of eliciting the response in question. In the pilot study cards 7GF, 6BM, 7BM, and 2 accomplished this very well. The remaining cards were arranged in descending order of their probability of eliciting this response. These latter cards were included in order to test the effect of reinforcement on cards which would not typically elicit the parent-child interaction response.

¹This study is based on a Master's thesis done by the junior author under the direction of the senior author.

²Now at the University of Tennessee.

Two groups of 20 subjects each were employed. Subjects were assigned randomly. Group I, the experimental group, consisted of 12 males and 8 females. Group II, the control group, included 13 males and 7 females. Prior to the presentation of the cards the following instructions were read to each subject:

I have a number of cards with pictures on them. I would like for you to make up a story to fit each picture. In your story I would like for you to tell me what events led up to the event shown in the picture, what is happening in the picture, how the characters feel, and what is going to happen next. I am interested in what kind of stories college students tell, so obviously there are no right or wrong answers or stories. The story you tell is entirely up to you. But remember, make sure you tell me what events led up to the event shown in the picture, what is happening in the picture, how the characters feel, and what will happen next.

Before card 16 was presented these additional instructions were given: "The next card is blank so you will have to make up the situation, the characters, and the story." Each subjects' stories were then recorded on a tape recorder.

Each time a subject in Group I gave a parent-child interaction response, the experimenter responded with an "mmm-hmm". The subjects in Group II were allowed to relate their stories without any response from the experimenter. In order for a given response to be considered one indicating parent-child interaction and thus one to be reinforced, one of the following criteria had to be met: (1) The response had to indicate interaction between the parent and child, for example, a mother talking to her daughter or a father advising his son; or (2) the response had to express an emotional interaction between the parent and child, for example, a mother is sad because her son is leaving home, a son is happy because his father has made him a partner in his business; or (3) the response had to indicate an evaluation of one by the other, for example,

a mother disapproves of her son leaving home, a father is critical of his son. Once the characters were identified as a parent and a child the use of pronouns sufficed to identify them. Simply identifying characters in the pictures as a parent or as a child did not warrant a reinforcement.

Immediately after the conclusion of the session, each subject in the experimental group was informally interviewed to determine if he had been aware of the real purpose of the experiment. None of the subjects reported such awareness.

RESULTS

Prior to any test of the significance of the difference between the two groups on the experimental hypothesis, it was necessary to demonstrate that the reinforcements given to Group I did not increase the length of the stories obtained from this group (reinforcement of speaking), which due to the longer stories would increase the probability of occurrence of parent-child responses. This was done by taking a total word count for each group and then testing for the significance of the difference between the two groups. The total number of words for Group I was 19,463 (Mean Ranks = 19.9) and for Group II 21,240 (Mean Ranks = 21.1). The Mann-Whitney U for the significance of the difference was 188, which is not significant at the 5% level ($U = 127$ for a two-tailed test at the 5% level). Neither was there a significant difference in total words between the females of the two groups (Group I: total words = 7702, Mean Ranks = 6.25; Group II: total words = 7531, Mean Ranks = 8.57; $U = 32$).

The tapes on which the stories were recorded were replayed and the number of reinforcements given Group I were counted. Since this merely involved a count of the number of "mmm-hmms" in each story, it was felt that it was unnecessary to do a reliability check. In the case of Group II, however, both authors lis-

tened to the tapes independently and made their counts on the basis of where the "mmm-hmm" would have been given had these subjects been reinforced. This procedure provided a reliability check on the adequacy of the criteria for reinforcement outlined above. One author's count was 200 and the other 198. A correlation of the two yielded a Pearson r of .97.

The Mann-Whitney U was employed to analyze the difference between the groups. Tests of significance were run between the total number of themes for the two groups, and then between the total number of themes for the first four cards for the two groups. This latter analysis was done to see what effect reinforcement had on cards where the "pull" was such that there was a high probability of obtaining parent-child interaction responses. The same procedure was followed for the total number of themes for the last six cards. This was carried out to measure the effect of reinforcement on cards which typically would not elicit the parent-child interaction theme. Post hoc analyses were then done on the intra- and inter-sex differences on the total number of themes, total themes on the first four cards, and for the last six

TABLE I—Mann-Whitney U Comparison of Groups I & II^a

Variable	Total number of themes		Mean Ranks	U	p^a
	Gr. I	Gr. II			
All cards	257	208	23.27	144.5	>.05
First four cards	180	134	22.25	123.5	<.05
Last six cards	77	66	24.32	165	>.05
			16.67		

^aN of each group = 20

^bTwo tailed test

cards. These results are reported in Tables I and II.

As is readily apparent in Table I, the only significant difference between Groups I and II was on the total number of themes for the first four cards. A further breakdown of this along sex lines (see Table II) indicates a significant difference between females in the two groups for total number of themes and also for the total number of themes for the first four cards and for the last six cards.

DISCUSSION

The only significant difference observed between the two groups was on the total number of themes given to the first four cards. The differences on the total themes given to all cards and to the last six cards, while in the

TABLE II—Mann-Whitney U Comparison of Males and Females Between Groups

Variable	Total Number of themes	Mean Ranks	U	p^a
Females, Groups I (N = 8) & II (N = 7), total cards	Gr. I 94	9.19		
	Gr. II 60	6.64	18.5	.308
	Gr. I 163	14.62		
Males, Groups I (N = 12) & II (N = 13), total cards	Gr. II 140	11.50	58.5	>.05
	Gr. I 69	9.75		
Females, first four cards	Gr. II 40	6.00	14	.12
	Gr. I 111	15.30		
Males, first four cards	Gr. II 94	10.88	50.5	>.05
	Gr. I 25	8.94		
Females, last six cards	Gr. II 20	6.93	20.5	.43
	Gr. I 52	14.17		
Males, last six cards	Gr. II 46	11.92	64	>.05

^aTwo tailed test

predicted direction, did not reach significance. This would seem to indicate that on those cards having a high "pull" for the parent-child interaction theme, the verbal reinforcements of an examiner can serve to increase the frequency of such a theme beyond that which might be expected under neutral conditions. The fact that significant differences on the last six cards could not be demonstrated suggests that the effect of reinforcement is negligible unless the "pull" of the card is appropriate to the theme being reinforced.

A post hoc analysis of sex differences reveals that there were significant differences between females in the two groups on all three comparisons: total cards, first four cards, and last six cards. Due to the small *N*s (7 and 8) involved, such differences must be interpreted with caution. These results also suggest that females accounted for most of the difference in the major finding of the study. The most obvious interpretation of such results is that females were more susceptible to reinforcement than were males. Such a conclusion has support from the results of other studies, indicating that females tend to be more conforming (Beloff, 1958), that they are more persuasible (Janis & Field, 1958), and that they tend to be more susceptible to the pressure of group norms (Tudenhams, 1958).

An inspection of Table II reveals that there were no significant differences between males in the two groups on any of the three breakdowns of the cards. Consequently, it would appear that males give about the same number of parent-child interaction themes regardless of conditions. An inspection of Table III indicates that in both groups and in all three breakdowns, males received more reinforcements than did females. Though none of the differences were significant, which may be due to the small *N*s involved, the possibility does arise that males simply

give more of the themes in question, reaching an asymptote which the effect of reinforcement could not overcome. Females on the other hand, giving fewer parent-child interaction responses and not reaching an asymptote, were thus more subject to the effect of reinforcement.

TABLE III—Mann-Whitney U Comparison of Males & Females within Groups

Variable	Total Number of themes	Mean Ranks	U*
Group I			
Males (n=12) vs females (n=8), total cards	163 94	10.8 10.0	44
Males vs females, first four cards	111 69	10.92 9.87	43
Males vs females, last six cards	52 25	11.17 9.50	40
Group II			
Males (n=13) vs females (n=7), total cards	140 60	11.42 8.79	33.5
Males vs females, first four cards	94 40	11.35 8.93	34.5
Males vs females, last six cards	46 20	11.46 8.71	33

*None of the U's was significant at the .05 level (two tailed test).

Finally, it seems plausible to account for the lack of more significant differences between the two groups on the basis of the fact that the response unit employed in this study was of too great a length for the reinforcement to effect it in its entirety. Thus it is quite possible that something other than the entire parent-child interaction theme (the last word spoken prior to reinforcement, for example) was being influenced by the reinforcement, the effect of which did not become apparent in the analyses carried out in this study. Further studies are now being planned to investigate this possibility.

CONCLUSIONS

The results of this study suggest that TAT examiners can influence

the output of themes by their verbal behavior. This seems to be especially true in the case of female subjects. Thus users of the TAT should be alert to the fact that they may inadvertently increase the frequency of a given theme. In view of the role the "pull" of the card plays in increasing such frequencies, caution should be exercised in interpreting themes from subjects for whom the examiner has selected specific cards in an effort to obtain information to validate a clinical hypothesis that he has formed about the subject.

REFERENCES

- Bellak, L., The thematic apperception test in clinical use. In *Projective psychology*, Abt. L., & Bellak, L., (eds.), New York: Grove Press, 1959.
- Beloff, Halla. Two forms of social conformity: Acquiescence^a and conventionality. *J. abnorm. soc. Psychol.* 1958, 56, 99-104.
- Greenspoon, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, 68, 409-416.
- Gross, L. Effects of verbal and nonverbal reinforcement in the Rorschach. *J. consult. Psychol.*, 1959, 23, 66-68.
- Janis, I. L., & Field, P. B. Sex differences and personality factors related to persuasibility. In C. I. Hovland & I. L. Janis (eds.), *Personality and persuasibility*. New Haven: Yale University Press, 1959, 55-68.
- Krasner, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, 55, 148-170.
- Mandler, G., & Kaplan, W. K. Subjective evaluation and reinforcing effect of verbal stimulus. *Science*, 1956, 124, 582-583.
- Murray, H. A. Manual for the thematic apperception test, Cambridge: Harvard University Press, 1943.
- Rotter, J. B. Thematic apperception tests: Suggestions for administration and interpretation. *J. Pers.*, 1946, 15, 70-92.
- Taffel, C. Anxiety and the conditioning of verbal behavior. *J. abnorm. soc. Psychol.*, 1955, 51, 496-501.
- Tatz, S. J. Symbolic mediation in "learning without awareness". Paper read at East. Psychol. Assoc., Atlantic City, March, 1956.
- Tuddenham, R. D. The influence of a distorted group norm upon individual judgment. *J. Psychol.*, 1958, 46, 227-241.
- Wickes, T. A. Examiner influence in a test situation. *J. consult. Psychol.*, 1956, 20, 23-26.
- Wilson, W. C., & Verplanck, W. S. Some observations on the reinforcement of verbal operants. *Amer. J. Psychol.*, 1956, 69, 448-451.

Received January 17, 1962

A Study of Some Communicable Measures for the Evaluation of Human Figure Drawings¹

DEODANDUS J. W. STRUMPFER² AND ROBERT C. NICHOLS³
Purdue University

In the Administration of the Draw-a-Person Test (DAP) the testing situation is only very superficially structured by non-specific instructions. Since the test variables are so little controlled with respect to their number and selection, evaluation of a DAP protocol usually requires highly complex judgments, the bases of which are not readily communicable.

Most frequently the DAP is evaluated in either a molecular or a molar fashion. Machover's (1949) approach consists of a molecular evaluation of numerous specific graphic and content details, about which she has formulated interpretive hypotheses. Swensen (1957) reviewed the research literature on these hypotheses and concluded that more of the evidence is in contradiction than in support of them. Subsequent to her monograph, Machover (Witkin *et al.*, 1954) developed a check-list of drawing items that discriminated between groups of subjects (Ss) achieving high, intermediate and low scores on tasks requiring spatial judgments in very unstructured situations. Although Machover's attempts still require quite subjective and complex judgments, the appealing aspect of a molecular approach is that it holds the possibility of achieving communicability. Formidable problems are the multiplicity

of cues and the integration of judgments.

In a molar approach to the DAP an attempt is made to evaluate the structure and the content globally and impressionistically. It has frequently been stated (e.g. Dunn & Lorge, 1954), possibly with gestalt reasoning, that this is the more fruitful approach. It is, however, obvious that even less communicability is achieved by a molar approach than by the Machover approach. The former is by definition unformulated, insightful and intuitive. Research findings on molar approaches showed them to be of questionable validity too. Whitmyre (1953) and Sherman (1958b) both demonstrated a high degree of relationship between judgments about the artistic quality of human figure drawings and judgments about the level of adjustment based upon the drawings. It would appear that the material to be judged is so complex that only its most prominent aspect can be handled adequately. In neither of these studies did either the art or the adjustment ratings show consistently significant relation to the dichotomy of psychiatric vs. non-psychiatric Ss. Smith (1953) found that clinicians could not differentiate better than chance between drawings by schizophrenic and nonschizophrenic Ss. In a study by Wallon (1959) four out of five judges obtained a number of correct identifications of drawings by normals, neurotics and psychotics which exceeded chance expectations. The diagnostic ratings of three of these judges were, however, significantly related to the age of the Ss. This suggests the possibility that the fact that the three groups differed significantly in mean age may have con-

¹This paper is based upon a Ph.D. thesis submitted to the faculty of Purdue University (1959) by the first author, working with the second author as major professor. In the initial stages this investigation was supported by a summer research grant from the Purdue Research Foundation.

²Now at the Department of Psychology, Potchefstroom University for Christian Higher Education, Potchefstroom, Union of South Africa.

³Now at National Merit Scholarship Corp. Evanston, Illinois.

tributed to the successful discrimination.

The important relationship between validity of judgment and the degree of complexity of the material judged was pointed out by Hamlin (1954) in a discussion of the implications of ten studies concerning judgments about projective and other clinical materials. In general, negative results were found when the units of material judged were either too simple or molecular, or too complex to be manageable within the limits of the particular experiment. In three of these studies (Albee & Hamlin, 1949, 1950; Bialick & Hamlin, 1954), however, a product scale consisting of samples of the projective material under consideration was first established. Significant correlations were found between outside criteria and judgments with the aid of these scales. Essentially simplification of a molar approach was involved here. The task of the judge became one of matching the new projective material with the illustrative materials of the scale. Such judgments are limited: they are neither overly simple nor overly complex. Furthermore, the bases for the judgments can be made quite communicable.

In the present study a number of DAP scales that constitute attempts to make the bases for judgment communicable were employed. Five of these scales were product scales and two required relatively objective molecular judgments. Two operational measures, based on the height of the drawn figures, were also included.

The first objective of the investigation was to determine, to some extent, the psychological meaning of these measures. In a discussion of the criterion problem in personality research Thurstone (1955) proposed a procedure of obtaining an intermediate criterion by comparing scores on a new test with each of a large number of responses on a personality self-appraisal. By this procedure he hoped to judge whether a new test measured an aspect of personality at all, and

also to ascertain the traits that are indicated by the new test. The suggested procedure was to determine which of the self-appraisal items were significantly related to the new test score. Those items would then be examined in an attempt to ascertain whether there was any "psychological unity" among them. Thurstone commented that this is clearly an inductive procedure that depends on the ability of the investigator to discover such unity among the self-appraisal items found to be significantly related to the test. In the present study this procedure was followed, using the California Psychological Inventory (CPI; Gough, 1957) as a pool of 480 items that cover a wide area of self-appraisal.

The second objective was to determine whether the DAP measures were significantly related to actual levels of personal adjustment, in terms of the broad categories of normal, neurotic and schizophrenic adjustment.

PROCEDURE

Subjects

Group 1 consisted of male students in an introductory psychology course at Purdue University. They were tested in two groups and both the DAP and the CPI were administered. The tests were presented as part of a research project during regular class meetings. A special test booklet was used for the DAP, containing instructions, blank pages (8½ x 11 inches) for the drawings, and a sentence completion test. Ss were instructed to start working on the sentence completion test upon completion of the first drawing. After everyone had completed the first drawing, the group started the second drawing and proceeded to the rest of the sentence completion test until everyone was finished with the second drawing. The sentence completion test served to prevent Ss who had completed a drawing from disturbing others. All Ss were provided with medium-soft (No. 2) pencils with erasers. A number of Ss were excluded because they did not draw full-

length figures or because they had received drawing instruction beyond that generally given in grade school. DAP's were thus available for a group of 107, and 91 of these had also taken the CPI. The mean age of the entire group was 20.06 ($SD = 4.07$). They ranged from freshmen to seniors, with the mean number of years of college work completed being .76 ($SD = .68$).

For the purpose of cross-validation of the findings on Group 1, the DAP and the CPI were administered to Group 2, consisting of 50 male students at Butler University.⁴ These students were enrolled for courses in introductory psychology ($n = 30$) and educational psychology ($n = 20$). Their mean age was 20.92 years ($SD = 2.85$). The group included students in all four years of college with a mean of .98 years of college work completed ($SD = .55$). The differences between the means for age and educational achievement of Groups 1 and 2 were not significant. The difference in variance for age was significant (.01 point) but not the difference in variance for educational achievement.

Group 3 consisted of 90 veterans of the U.S. Armed Forces.⁵ Three subgroups were included: 30 "normals", 30 neurotics and 30 schizophrenics. The normal Ss were medical and surgical patients in the VA Hospital at Indianapolis, Indiana, who were tested just prior to their discharge from the hospital. They were selected after consultation with the resident physicians or surgeons and the nursing staff of the wards on which they were patients. None of them had anamnestic indications of a psychiatric condition. No patients with psy-

chomatic conditions were included, and amputees and paraplegics were also excluded.

DAP's for the neurotic subgroup were obtained from the files of the neuro-psychiatric services of the VA Hospital in Indianapolis ($n = 10$) and the VA Mental Hygiene Clinic in Detroit, Michigan ($n = 20$). The Indianapolis patients were in-patients at the time of testing, while the Detroit patients were outpatients. The psychiatric diagnosis was anxiety neurosis in the majority of cases, with a number of additional cases of mixed neurosis, conversion reaction, dissociative reaction, obsessive-compulsive neurosis and hypochondriasis.

The DAP's for the schizophrenic Ss were obtained from the files of the neuropsychiatric services of the VA Hospitals in Indianapolis ($n = 15$) and in Ann Arbor, Michigan ($n = 15$). All of these patients were hospitalized on the psychiatric wards of the two hospitals at the time of testing. The majority of them were diagnosed as either chronic undifferentiated or paranoid schizophrenia, with a number of additional catatonic, hebephrenic and mixed cases. Cases with an implication of organic brain damage were excluded. A limiting factor about this sub-group was the fact that they were probably not fully representative of a schizophrenic population. The most disturbed patients frequently will not cooperate in a figure drawing task and such persons were thus excluded by the manner of selection of this sample.

The neurotic and schizophrenic cases were selected so as to match the normal sample on a group basis with respect to mean age and mean number of years of formal education. Table I shows the ranges, means and standard deviations for age and educational achievement for these three subgroups. Tests for the significance of differences between the means and for homogeneity of variance failed to show significant differences between the groups.

⁴The authors wish to thank Drs. J. William Hepler, Roger Coulson and David Torbett for making their classes available for testing.

⁵The authors wish to thank Drs. Ruth Baker Hines and Robert A. Wagoner of the VA Hospital, Indianapolis, Ind., and Dr. Edward J. Wallon of the VA Mental Hygiene Clinic, Detroit, Mich., for their assistance in connection with the collection of test data.

All drawings of Group 3 were done on blank sheets of paper, 8 x 10½ inches in size. All manner of identification was removed from all of the drawings and code numbers were assigned at random to the Ss in the subgroups. An A or a B appeared with the code number on each drawing to indicate the sequence of drawing. In order to give a "had-been-filed" appearance to the drawings of the normal Ss too, these were stapled together and the staples removed again.

Drawing Variables

Brief descriptions of the measures that were used are given below. The literature about these measures has been reviewed by Strumpfer (1959). The following product scales were used:

(1) The *Artistic Quality Scale* was developed by Wagner and Schubert (1955) to quantify global judgments about the artistic quality of DAP's by late adolescents and young adults. Their Ss were "normal, average, reasonably adjusted, ordinary people" (p. 1), although somewhat selected in terms of educational achievement. Four series of seven illustrative drawings each were selected by unanimous agreement of three judges: front and profile series for both male and female figures. Verbal descriptions of the scale points were added to aid in ratings, as well as instructions in connection with the rating of unusual drawings. The verbal scale was extended by one point at each end to include extremely superior and inferior ratings. The scale for male figures by males is used to rate male figures by female Ss, and the scale for female figures by females to rate female figures drawn by male Ss. In the present

study the male and female figures drawn by each S were rated separately by each judge and these two ratings were summed to obtain the Artistic Quality Score for the S.

(2) The *Adjustment Scale* was developed by Albee and Hamlin (1949) to quantify global judgments about the level of adjustment reflected in the DAP. Adjustment was defined as, "what is usually denoted by that term—psychosexual maturity, ability to get along with other people, and perception of reality in a way that agrees with our cultural standards" (p. 390). Drawings were obtained from ten Ss, carefully selected to represent a wide range of adjustment, from a normal S to an hallucinated, delusional psychotic. These drawings were presented to 15 judges for paired comparisons about the level of adjustment reflected in them. Mean scale values were thus derived, placing the drawings on a continuum with unequal intervals. The product scale consists of ten pairs of drawings (male and female figures together). In matching a new drawing with those of the scale, the two drawings by the S have to be considered together, so that a composite rating of both drawings is done. In the present study the ranks assigned to the ten pairs of drawings by Albee and Hamlin's judges were used as "scores", instead of the scale values employed by the authors of the scale. The rationale for this procedure was that the rank of a pair of illustrative drawings actually constitutes a more appropriate "score" when a new pair is matched with it; the matching procedure is not accurate enough to warrant the assumption that the matched pair falls at exactly the same point along the continuum as the illustra-

TABLE I—Ranges, Means and Standard Deviations for Age and Educational Achievement of Three Subgroups of the VA Sample

Groups	Age			Education		
	Range	Mean	SD	Range	Mean	SD
Normal	25-43	33.63	5.81	6-16	10.20	2.55
Neurotic	25-42	33.67	5.19	6-16	10.00	2.59
Schizophrenic	25-44	33.40	5.58	6-16	10.27	2.41

tive pair, and, therefore, the scale value of the latter cannot be assigned to the former.

(3) The *Sexual Differentiation Scale* was developed by Swensen (1955) to quantify judgments about the degree to which the sexual characteristics of the male and female drawings were differentiated by the S. The assumption was that adequate sexual differentiation on the DAP indicates adequate sexual identification on the part of the S. Impairment might reflect a general deficit, as in the case of psychosis, or an impairment confined mainly to the sexual area, as in the case of homosexuality. DAP's from the files of a VA mental hygiene clinic were used. A five-point-scale was constructed on the basis of sortings by seven judges, with three pairs of drawings (male and female) to illustrate each point. Verbal descriptions of these points were added. The scale was expanded with four points to be used for rating drawings that appear to fall in between the illustrated points. In matching new drawings with those of the scale, the two drawings by the S have to be considered together, so that a composite rating of both drawings is done.

(4) The *Maturity Scale* was developed by Dunn and Lorge (1954) to quantify judgments about the maturity of a figure drawing as a total representation. Drawings by normal and deviant children and adolescents, ranging in age from five to 20, were ranked by four judges. Every fifth percent position was selected in the series for each sex, resulting in two scales of 20 illustrative drawings each. In the present study the procedure of the Artistic Quality Scale was followed, rating the female figures by male Ss on the scale for female figures by females. Three scores were derived: a score for the male figure, a score for the female figure, and the sum of these two (Total Maturity Score).

(5) The *Aggression Scale* was developed by Strumpfer (1959) to quantify judgments about the amount of

aggressiveness or hostility depicted in a drawing. Separate scales were developed for male and female drawings. DAP's by 107 male college students (Group 1) in an introductory psychology course were rated by nine judges. On the basis of a study by Goldstein and Rawn (1957) of the effects of experimentally induced feelings of aggression, these judges were requested to keep the following points in mind while rating the drawings: slash-line mouth, detailed teeth, nostril emphasis, spiked fingers, clenched fist, squared shoulders and toes in a non-nude figure. To this list were added general facial expression and instruments of aggression. The judges were, however, requested to use any other cues that they use habitually in their clinical work. The male and female series of drawings were each sorted into five piles, using a forced normal distribution with the following percentages (suggested by Webster, 1956) in each pile: 9, 20, 42, 20, and 9 per cent. The mean and the standard deviation were calculated for the nine ratings of each drawing. The three drawings that showed a mean closest to the rating point under consideration, and the smallest standard deviation, were selected to illustrate that scale point. The scales for male and female drawings thus each consisted of 15 illustrative drawings. There was, however, a considerable amount of overlap among the ratings of drawings selected to illustrate adjacent scale points. In the subsequent use of the scale global judgments were emphasized but the above-mentioned details were considered as aids in rating difficult drawings. The Aggression score was the sum of the scores for the male and the female figures.

In addition to the above product scales, two point scales and two height measures were used.

(1) The *Body Image Disturbance Scale* was developed by Fisher (1959). It is an abbreviation of Machover's checklist (Witkin, et al., 1954) and

TABLE II Interrater and Rate-Rate Reliability Coefficients of DAP Scales for College (N = 107) and VA (N = 90) Samples

	College		Interrater VA	Rate VA
	r	r _{s-n}	r	r
Artistic Quality	.85	.92	.91	.95
Adjustment	.40	.57	.62	.76
Sex Differentiation	.68	.81	.70	.82
Maturity: Male figs.	.54	.70	.88	.94
Maturity: Female figs.	.70	.82	.86	.92
Total Maturity	.64	.78	.92	.96
Aggression: Male figs.	.82 ^a	.90	.40	.57
Aggression: Female figs.	.76 ^a	.86	.31	.48
Total Aggression			.42	.59
Body Image	.76	.86	.89	.91
Weighted Flaw Score			.95 ^b	.95
Weighted Good Score			.96 ^b	.95
Net Weighted Score			.96 ^b	.96
% Raw G Score			.95 ^b	.93 ^b

^a Based on ratings by eight judges.

^b n=50.

retains 14 points based on simple and relatively objective observations and judgments. Penalty points are assigned for the presence of these signs.

(2) The *Buck Scale* is the Person section of the quantitative scale for the House-Tree-Person Test, developed by Buck (1948) for the purpose of determining intellectual ability from such drawings. The scale consists of a checklist of items of detail, proportion and perspective with different penalty and credit values. The two figure drawings of each S were scored separately and the two sets of raw scores were then summed. The four scores: *Weighted Flaw Score*, *Weighted Good Score*, *New Weighted Score* and *Percentage of Raw G*, were determined in the way described by Buck. He provides tables for converting these scores to IQ's, when they are based on drawings of a house, a tree and a person. This could, of course, not be done with scores based on two person drawings.

(3) The *height* of each figure drawing was measured in millimeters, ignoring minor details extending above and below the figure. These measurements were combined in two ways: (a) the *ratio* of the height of the male figure to the height of the female figure and (b) the *sum* of the heights

of the male and the female figures.

Judging the Drawings

Doctoral students in clinical psychology served as judges.⁶ They had all completed a minimum of one year of graduate work beyond the master's degree and all of them had completed nine credit hours of course work on projective techniques. Their amount of practical experience varied. They were, however, not considered as experts in the use of the DAP. For all analyses the sum of two sets of ratings were used, except in the case of the Buck scores where scores assigned by one judge were used. The drawings were rerated or rescored after a period of four to six weeks.

The *reliabilities* of the various measures are shown in Table II. The correlations between ratings or scores assigned by pairs of psychologists on each of the DAP scales are presented for both Groups 1 and 3. The reliabilities of their combined ratings or scores, as determined by the Spearman-Brown formula, are also shown. In the case of Group 1 the coefficients for the Aggression Scale are based on ratings by eight judges, combined into two groups of four, without the use

⁶ The authors wish to thank Karl Beck, Herbert Lange, John Mendenhall and Abe Ritzenhouse for serving as judges.

of a product scale (i.e. as obtained at the time of constructing the scale). Rate-rater reliabilities for Group 3 are also shown in Table II. In the case of the Buck scores (Group 3) both the interscorer and score-rescore reliabilities are based on only 50 randomly selected DAP's (19 normal, 15 neurotic and 16 schizophrenics Ss).

The drawings of Group 3 were also rated on all of the scales, except the Buck scale, by a 19-year-old college sophomore whose background in psychology consisted of one introductory course. His instructions for rating the drawings were the same as those of the psychologists. Table III shows the correlations between the ratings by this judge and the combined ratings by pairs of psychologists.

TABLE III—Correlations Between Combined Judgments of Two Psychologists and Those of a Non-Psychologist for Drawings by VA Sample (N = 90)

DAP Scale	r
Artistic Quality	.92
Adjustment	.80
Sex Differentiation	.86
Maturity: Male figs.	.86
Maturity: Female figs.	.93
Total Maturity	.93
Aggression: Male figs.	.56
Aggression: Female figs.	.54
Total Aggression	.55
Body Image	.90

ANALYSES OF DATA

Item-analyses of the CPI

Item-analyses of the CPI, using the DAP measures as criteria for the selection of upper and lower groups, were carried out on the data of the 91 Ss of Group 1 for whom both DAP's and CPI's were available. The procedure followed in these item-analyses was recommended by Webster (1956), in which the tests were grouped into five categories according to each of the drawing scores. Response frequencies were counted for each item separately for each category and a D score was calculated for each item by weighting the response frequencies in the five

categories as -2, -1, 0, +1 and +2. The significance of the D scores was evaluated according to the formula given by Webster (1956).

The number of CPI items that showed a significant relationship to each DAP measure are shown in Table IV. Among the 480 items a significant relationship would be expected by chance in the case of 24 items at the .05 level, approximately five at the .01 level, and none at the .001 level. Thus a number of items well beyond chance expectations were found to be related to the following DAP measures: Adjustment Scale, Maturity Scale for male figures, Weighted Good Score, Net Weighted Score, Height M/F, and Height M+F. These findings were cross-validated.

TABLE IV—Number of CPI Items Significantly Related to Each DAP Measure

DAP Measure	Significance level			Total number
	.05	.01	.001	
Artistic Quality	17	4	0	21
Adjustment	28	9	1	38
Sex Differentiation	26	3	0	29
Maturity: Male figs.	25	8	0	33
Maturity: Female figs.	27	3	1	31
Aggression:				
Male figs.*	20	1	0	21
Aggression:				
Female figs.*	17	5	0	22
Body Image	13	4	0	17
Weighted Flaw Score	15	10	0	25
Weighted Good Score	30	5	0	35
Net Weighted Score	31	15	14	60
% Raw G Score	21	5	0	26
Height: Male/Female	26	7	13	46
Height: Male+Female	38	13	2	53

* Based on mean of nine sets of ratings without use of product scale.

In the first step of cross-validation those items that were found to be related to a DAP measure were considered as items in a new CPI scale. A scoring key was made up in each such instance and the CPI's of Group 2 were scored on these new scales. Scores on the DAP measure under consideration were obtained from the DAP's of Group 2. Scores on the DAP-CPI scale and on the DAP measure were

then correlated, using a one-tailed test for the significance of r . These correlation coefficients are presented in Table V. Only in the case of the Weighted Good Score was the correlation with the DAP-CPI scale significant.

TABLE V—Correlations between DAP Measures and DAP-CPI Scales for Cross-Validation Purposes ($N = 50$)

DAP Measure	r
Adjustment	-.064
Maturity: Male figs.	.000
Weighted Good Score	.241*
Net Weighted Score	.108
Height: Male/Female	.197
Height: Male+Female	-.030

*Significant at .05 point.

In this instance the second step of cross-validation was to repeat the item-analysis. Since a one-tailed test of significance was made, the .05 point was used in the calculation of D values in Webster's formula. Twenty of the 35 items that originally showed a relationship to the Weighted Good Score continued to show a significant relationship. This number is well beyond chance expectation. These 20 CPI items are shown in Table VI, with the responses in the direction that would be expected with a person obtaining a low Weighted Good Score.

For the inductive evaluation of these items, as suggested by Thurstone (1955), these items were presented to four experienced psychologists, two holding Ph.D.'s and two M.A.'s.

As an additional study of the inventory test correlates of the DAP measures, these were correlated with the 18 CPI scales (Gough, 1957). Approximately 13 correlation coefficients significant at the .05 level could be expected to occur by chance alone among the 252 that were computed. Actually only three reached this level of significance.

Comparisons of Normal, Neurotic and Schizophrenic Samples

A single classification analysis of variance was carried out for each DAP

measure to determine whether it differentiated significantly between the Ss in the normal, neurotic and schizophrenic subgroups of Group 3. Hartleys F_{\max} test. (Walker & Lev, 1953) was used to test for homogeneity of

TABLE VI—CPI Items Related to Weighted Good Score

No.	Item
<i>Self-consciousness vs. self-confidence</i>	
177.	I am certainly lacking in self-confidence. (T)
243.	I am often bothered by useless thoughts which keep running through my mind. (T)
279.	I often get disgusted with myself. (T)
320.	I would be willing to describe myself as a pretty "strong" personality. (F)
369.	I seem to do things that I regret more often than other people do. (T)
458.	People who seem unsure and uncertain about things make me feel uncomfortable. (T)
<i>Egotism vs. altruism</i>	
57.	I have sometimes stayed away from another person because I feared doing or saying something that I might regret afterwards. (T)
375.	There are certain people whom I dislike so much that I am inwardly pleased when they are catching it for something they have done. (T)
468.	I like to eat my meals quickly and not spend a lot of time at the table visiting and talking. (T)
<i>Withdrawal vs. outgoing interests</i>	
9.	I usually go to the movies more than once a week. (F)
140.	I enjoy hearing lectures on world affairs. (F)
228.	I like to read about history. (F)
244.	If I were a reporter I would like very much to report news of the theater. (F)
355.	I have strong political opinions. (F)
450.	I get sort of annoyed with writers who go out of their way to use strange and unusual words. (T)
<i>Inadequacy vs. adequacy of identification with moral standards</i>	
42.	I sometimes pretend to know more than I really do. (T)
47.	Women should not be allowed to drink in cocktail bars. (F)
96.	I take a rather serious attitude toward ethical and moral issues. (F)
155.	A person should adapt his ideas and his behavior to the group that happens to be with him at the time. (T)
<i>Miscellaneous</i>	
19.	I think I would like the work of a building contractor. (T)

variance. The F values obtained in the analysis of variance and the F_{\max} values are shown in Table VII. None of them was significant at the .05 point.

TABLE VII—Results of Analyses of Variance on DAP Measures for Normal, Neurotic and Schizophrenic Samples

DAP Measure	F	F_{\max}
Artistic Quality	.45	1.55
Adjustment	.05	1.31
Sex Differentiation	1.88	1.40
Maturity: Male figs.	.19	1.11
Maturity: Female figs.	.38	1.63
Total Maturity	.29	1.31
Aggression: Male figs.	.33	1.54
Aggression: Female figs.	.17	1.44
Total Aggression	2.02	2.33
Body Image	.02	1.39
Weighted Flaw Score	.13	1.26
Weighted Good Score	.90	1.71
Net Weighted Score	.46	1.66
% Raw G Score	.69	1.92
Height: Male/Female	1.40	2.05
Height: Male+Female	.16	1.24

For $df = (2, 87)$, $F = 3.10$ at .05 point

For $df = (3, 29)$, $F_{\max} = 2.36$ at .05 point

DISCUSSION

Reliability

The interrater reliabilities reported in Table II compare favorably with those reported previously for the same scales when used on similar or different groups (Wagner & Schubert, 1955; Albee & Hamlin, 1950, 1950; Swensen, 1955; Cutter, 1956; Dunn & Lorge, 1954). In general, the reliability of these scales is not high enough to warrant their use in individual predictions but the majority of them could be used with a degree of consistency satisfactory for group-screening purposes. The overall impression is that an acceptable degree of reliability of judgment can be achieved when the units of judgment are made simple enough to be communicable. Both interrater and rate-rater reliabilities tended to increase in inverse proportion to the complexity of the units to be judged. Factors of importance in this connection seem to be whether the concept on which the judgments are based is explicit or not; whether a composite judgment of the two figure

drawings is made or whether they are judged singly; whether the scale is well-illustrated, either by having more than one illustrative drawing at each scale point or by having a long enough scale; and whether explicit verbal descriptions of the drawings at each point of the scale are provided.

The exceptionally low reliability of the Aggression Scale probably results largely from the amount of overlap between the characteristics of the illustrative drawings at the different scale points. Another factor is the narrow range of only five points on the scale.

The fact that the interrater reliabilities of the VA sample are consistently higher than those of the student sample could be related to a wider range of talent in the former sample. Some of the drawings of the VA sample were executed equally as well as the best of the student drawings, while others were extremely poor.

The correlations between the combined judgments of a pair of psychologists and those of a non-psychologist are roughly comparable to the interrater reliabilities (Tables II and III). They indicate that the task of rating is simple enough on the DAP measures concerned that no extensive training is needed for their use. (The fact that these correlations tended to be higher than the interrater reliabilities for two psychologists probably resulted from the more stable criterion of combined judgments.) These findings are in agreement with those of Albee and Hamlin (1950) and Wagner and Schubert (1955) that optimally simplified judgments could be made equally well by clinicians and clinically unsophisticated judges. Potentially the product scale approach thus seems to be valuable for large scale screening procedures employing relatively untrained personnel.

Item-analyses of the CPI

In the empirical construction of the CPI scales item-analyses were performed, using diverse criteria, and the

self-appraisal items were in this way shown to have some validity for the description of personality characteristics. The results of the item-analyses in the present study (Tables IV and V) show that, with one exception, the DAP measures were not related to a significant number of these items. According to Thurston's (1955) reasoning, it thus seems doubtful that these DAP measures are influenced to a significant extent by personality factors. The fact that positive results were obtained in the case of one of the DAP measures tends to make the negative results on the other measures even more meaningful. The nonsignificant correlations between the DAP measures and the CPI scales also tend to emphasize the lack of relationship between the DAP variables and personality characteristics.

In an inductive evaluation of the CPI items that showed relationship to the Weighted Good Score (Table VI) it has to be kept in mind that the inferences cannot be generalized to groups markedly different from the student sample on which these findings were obtained. It should also be kept in mind that the CPI responses reflect the way in which these Ss see and describe themselves and that the responses are not necessarily literally related to actual behavior.

The interpretations of the items in Table VI by four psychologists showed enough agreement to warrant the conclusion that these items have some aspects in common. Only one item (no. 19) does not have anything in common with the remaining 19. The following traits seem to be described as indicated by grouping of the items in Table VI:

Self-consciousness vs. self-confidence

Egotism vs. altruism

Withdrawal vs. outgoing interests

Inadequacy vs. adequacy of identification with moral standards

These traits can be conceived of as being interrelated and interdependent, so that "psychological unity" (Thurstone, 1955) seems to be pres-

ent. The Weighted Good Score thus seems to be, not only a measure of intellectual functioning, but also related to some of the personality characteristics that Ss ascribe to themselves. The empirical meaning of these self-descriptions remains to be determined.

Comparisons of Normal, Neurotic and Schizophrenic Samples

The personality characteristics that are purportedly measured by the different DAP scales could be conceived as subject to the changes observed from normality to neurotic or schizophrenic adjustment. The data presented in Table VII, however, show that not a single one of the 16 DAP measures was able to differentiate significantly between the three broad levels of adjustment represented in the VA sample. Thus the relationship between whatever these drawing variables do represent and these levels of adjustment is seriously questioned.

The findings on the Artistic Quality Scale are in agreement with those of Whitmyre (1953) and Sherman (1958 a and b), neither of whom, however, used product scales. The findings on the Adjustment and Sexual Differentiation Scale are contrary to the reports of Albee and Hamlin (1950) and Swensen (1955). In both instances the scale concerned was found to have discriminative value for such broad levels of adjustment. The findings of Albee and Hamlin have not been cross-validated but cross-validations by Sherman (1958) and Cutter (1956) left serious doubt about the validity of the Sexual Differentiation Scale for purposes of discriminating between levels of general adjustment, as well as sexual adjustment in particular. The Maturity and Body Image Disturbance Scales have not been used in research of this kind. Buck (1948) suggested that the different scores on his scale are influenced differentially by psychopathology. This is not borne out by the group comparisons in the present study.

With regard to the negative findings on the height variables in this investigation, it should be recalled that Reed (1957) found a significant difference between normal and psychotic Ss with respect to the relative height of male and female figures but that the relationship did not hold up on cross-validation. Hoyt and Baron (1959) found a significant relationship between the size of same-sex drawings and high and low Manifest Anxiety Scale scores in a group of psychiatric patients.

SUMMARY

Various relatively communicable measures for the evaluation of human figure drawings, including five product scale, two point scales and two height variables, were investigated as to: (1) their psychological meaning, by means of item-analyses on the California Psychological Inventory; and (2) their relationship to actual level of personal adjustment in terms of the broad categories of normal, neurotic and schizophrenic adjustment. Subjects were two groups of college students and a group of VA patients.

Among the 14 items-analyses of the psychological inventory, with the upper and lower groups selected on the basis of ratings or scores on each of the drawing measures, only one measure showed a significant number of the items to be related to it after cross-validation. This was interpreted as indicative of questionable validity for the drawing measures, since the self-appraisal items have previously been shown to be related to personality characteristics.

Analyses of variance of the data of normal, neurotic and schizophrenic groups failed to show significant differences between the means or variances of the three groups on any of the measures. The validity of these measures for the diagnosis of personality disorders was thus considered as doubtful.

REFERENCES

- Albee, G. W. & Hamlin, R. An investigation of the reliability and validity of judgments inferred from drawings. *J. clin. Psychol.*, 1949, 5, 389-392.
- Albee, G. W. & Hamlin, R. M. Judgment of adjustment from drawings: the applicability of rating scale methods. *J. clin. Psychol.*, 1950, 6, 363-365.
- Bialick, I. & Hamlin, R. M. The clinician as a judge: details of procedure in judging projective material. *J. consult. Psychol.*, 1954, 18, 239-242.
- Buck, J. N. The H-T-P technique: a qualitative and quantitative scoring manual. *J. clin. Psychol.*, 1948, Monograph Supplement, No. 5.
- Cutter, F. Sexual differentiation in figure drawings and overt deviation. *J. clin. Psychol.*, 1956, 12, 369-372.
- Dunn, M. & Lorge, I. A gestalt scale for the appraisal of human figure drawings. *Amer. Psychologist*, 1954, 9, 357. (Abstract)
- Fisher, S. Body reactivity gradients and figure drawing variables. *J. consult. Psychol.*, 1959, 23, 54-59.
- Goldstein, A. P. & Rawn, M. L. The validity of interpretive signs of aggression in the drawing of the human figure. *J. clin. Psychol.*, 1957, 13, 169-171.
- Cough, H. G. *Manual for the California Psychological Inventory*. Palo Alto, Calif.: Consulting Psychologists Press, 1957.
- Hamlin, R. M. The clinician as a judge: implications of a series of studies. *J. consult. Psychol.*, 1954, 18, 233-238.
- Hoyt, T. E. & Baron, M. R. Anxiety indices in same-sex drawings of psychiatric patients with high and low MAS scores. *J. consult. Psychol.*, 1959, 23, 448-452.
- Machover, Karen. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Thomas, 1949.
- Reed, M. R. The masculinity-femininity dimension in normal and psychotic subjects. *J. abnorm. Psychol.*, 1957, 55, 289-294.
- Sherman, L. J. Sexual differentiation or artistic ability? *J. clin. Psychol.*, 1958, 14, 170-171. (a)
- Sherman, L. J. The influence of artistic ability on judgments of patient and non-patient status from human figure drawings. *J. proj. Tech.*, 1958, 22, 338-340. (b)
- Smith, Elgie. A study of sex differentiation in drawings and verbalizations of schizophrenics. *J. clin. Psychol.*, 1953, 9, 183-185.
- Strumpher, D. J. W. A study of some communicable measures for the evaluation of human figure drawings. Unpublished Ph.D. thesis, Purdue Univer., 1959.
- Swensen, C. H. Sexual differentiation on the Draw-A-Person Test. *J. clin. Psychol.*, 1955, 11, 37-40.
- Swensen, C. H. Empirical evaluations of human figure drawings. *Psychol. Bull.*, 1957, 54, 431-466.
- Thurstone, L. L. The criterion problem in personality research. *Educ. & Psychol. Measmt.*, 1955, 15, 353-361.

Albee, G. W. & Hamlin, R. An investigation

- Wagner, Mazie E. & Schubert, H. J. P. *D.A.P. quality scale for late adolescents and young adults*. Kenmore, N. J.: Delaware Letter Shop, 1955.
- Walker, Helen M. & Lev, J. *Statistical inference*. New York: Holt, 1953.
- Wallon, E. J. A study of criteria used to differentiate the human-figure drawings of normals, neurotics, and psychotics. Unpublished Ph.D. thesis, Purdue Univer., 1959.
- Webster, H. Transformed statistics for use in test construction. *Psychol. Bull.*, 1956, 53, 488-492.
- Whitmyre, J. W. The significance of artistic excellence in the judgment of adjustment inferred from human figure drawings. *J consult. Psychol.*, 1953, 17, 421-424.
- Wiikin, H. A., Lewis, H. B., Hertzman, M., Machover, Karen; Meissner, B. P. & Wapner, S. *Personality through perception*. New York: Harper, 1954.

Received November 25, 1961

The Magnitude of Four Experimental Needs as Expressed by Two Projective Techniques

ERIC C. THEINER
Lackland Air Force Base, Texas

Psychologists have long pondered the question of the effectiveness of projective techniques in eliciting covert personality dynamics. Despite the fact that research findings are often ambiguous and, at times, even contradictory, they suggest that the validity of projective measures is accepted so widely that there must, in fact, be some truth in the belief. That much of the scoring is highly subjective, and that often there appears to be little relationship between two separate measures (Dorken, 1953) has not caused any marked degree of consternation. There is, occasionally, a call to return to a firmer basis of belief. Windle (1952), for one, implied that the prognostic use of projective methods was not yet fully justified in terms of research in this area; cross-validation studies should be emphasized.

Criticism has been leveled particularly against the more structured of projective measures.

... many of the newer techniques, such as the Sentence Completion, and the Situation Completion appear so patently close to the conscious cognitive and conative levels of ego functions, that what is gained by objectivity, speed of administration and scoring is often offset by the shallowness, meagerness, and inconsequentiality of the material they convey (Kass, 1956, p. 269).

The same feeling was expressed by Holsopple and Miale who stated:

... the sentence completion tests have usually approached the individual too directly. They have asked for conscious report on problems, and therefore attacked and exposed the subject. The subject's conscious attitudes have masked the material from which valid inferences concerning more or less unconscious trends could be drawn (Holsopple *et al.*, 1954, p. 18).

Yet, the reputation of projective tests in general remains that they can elicit material largely preconscious or

unconscious. But since the degree of projection which a stimulus elicits is felt to be largely a function of its degree of ambiguity, heading tests under the broad label "Projective Techniques" is misleading. Any individual method is only as "projective" as the factual anchorage of the media of which it is comprised. Need patterns are operant to some degree at all times, and a projective test will probably reflect some of these needs. The expressed magnitude is a function of the degree of ambiguity of the stimulus, since the ambiguity influences the cognitive defensive processes which the projective media calls into play.

INTRODUCTION

This present investigation was carried out in an attempt to compare the frequency of expression of four manifest psychological needs as obtained from two projective techniques: the Thematic Apperception Test (TAT) and the Rotter Incomplete Sentences Blank (ISB). The purpose of these comparisons was to investigate the relationship of frequency of expression of certain psychological needs to the ambiguity of the specific projective media employed. Briefly, different tests were expected to produce significantly different results even on the same dimension, i.e., a specific need.

The rationale for this study proceeded from the assumption that different needs possess varying degrees of cultural desirability. It was assumed that most college Ss would try to emphasize favorable needs, and de-emphasize unfavorable needs in their behavior.

For present purposes, projection was broadly defined according to its most commonly accepted meaning;

i.e., the attachment of inner feelings and values to external phenomena. Each person introduces his own belief value matrix into the perceptions he holds of the outer world. With the reduction of context clues, the individual is forced to employ increasingly more of his personal value system; projection diminishing with an increase of structure in the stimulus.

Bellak defined the Thematic Apperception Test (TAT) as a "... technique for the investigation of personality as it manifests itself in interpersonal relations and in the apparent or manifest interpretation of the environment" (Bellak, 1959b, p. 185). The question of the psychological level of consciousness at which the TAT tests has remained largely moot. Murray (1948) leaned toward the view that the TAT was a rather deep indicator, disclosing those aspects of personality which were composed of repressed unconscious needs, and needs appearing in thought alone in undisguised form. Bellak (1959b) wrote that preconscious material of the type elicited by the TAT had unconscious determinants, while Balken and Vander Veer (1942) denied that the TAT tested these deeper layers of personality. A third interpretation advocated by Wyatt (1947) and Rosenzweig (1950) held that the level of expression shown on the TAT varied from time to time. This implied that content material in terms of ambiguity of stimuli was an irrelevant variable.

Rotter defined the Sentence Completion method as "... a semi-structured projective technique in which the subject is asked to complete a sentence for which the first word or words are supplied" (Rotter *et al.*, 1950, p. 3). He felt the ISB elicited information that the S was willing to give. The measure had some disguise of purpose, but not as much as certain other measures. Rhode stated that association operated in the method but that "... strong ego defenses and escape mechanisms are involved in

varying degrees while the subject is formulating his responses" (Rhode, 1957, p. 3). By and large, the majority of investigators have felt that while the ISB might have some heuristic value as a projective measure, it did not elicit material from the same depths of personality as did tests such as the TAT or the Rorschach.

In this review of the literature, only one specific research comparison between a sentence completion technique and tests such as the Rorschach and the TAT (Samuels, 1952) has been found. On the basis of that study, Samuels maintained that projective methods measure little in common, and that their value as personality assessors was apparently limited by low validity. The implication was that some projective tests do not measure more effectively at different personality levels from others.

Since a test such as the Incomplete Sentences Blank consists of a comparatively greater amount of structured material than is presented by either the Rorschach or the Thematic Apperception Test, it may be hypothesized that the ISB will produce results more in line with the S's own desired self-image, than will either of the latter two. The TAT, on the other hand, should bring forth a greater amount of latent content. Should divergencies then appear between these two forms, this is due not so much to a lack of consistency concerning the information disclosed, but rather because there are different rungs on the ladder of conscious control and, in this instance, less controlled levels are being investigated.

The writer believes that in the case of the average male college student, favorable needs include Achievement and Affiliation. On the other hand, the average college student will seek to minimize the role of such needs as Aggression and Succorance. For purposes of this present comparison, the actual conscious acceptability of the experimental needs was determined

by a trait rating scale given to the S's during the testing situation.

Acceptability or unacceptability of a need was defined by the effects which the individual felt it would have on his immediate social circle. While this is somewhat restrictive in itself, (being limited to family and friends), it is a measure of social desirability which is conducive to scientific scaling.

Another question is that of neutrality. It is obvious that an ISB statement such as "I hate . . .," or a TAT card showing hands grasping a struggling man will produce a higher percentage of aggressive responses than more innocuous stimuli. The objection that the different patterns which are elicited from the ISB as contrasted to the TAT are due to incorporate sets in the stimulus is answered by saying that if this were the case, a basic assumption in the use of projective techniques would be voided. The entire rationale of the method is contingent on the fact that these tests are sufficiently "setless" to force the S to create his own perceptual pattern, with a minimum employment of any sort of cue mechanisms.

A theoretical issue concerns the reason why it is expected that different results will be obtained from the two projective tests employed in this study. Cards 4 and 14 were selected as simple as the S's more ready personal association to a statement such as "I like . . ." than to a card with several figures, none resembling the S very closely. Consequently, with more direct association, there will be a greater effort to maintain the desired self-image. On the other hand, some theorists assert that different "levels of personality" are being tested (Bellak, 1959a, 1959b; Cattell, 1951; Eron *et al.*, 1948; Rosenzweig, 1950; Wyatt, 1947). This study has bypassed the issue in terms of test controls. Rather, the major hypothesis revolves solely about the significant difference in

need structure reflected by the two experimental measures.

PURPOSE

The purpose of this experiment is to investigate if a significant difference exists in the frequency of expression of needs Affiliation, Achievement, Aggression, and Succorance between the TAT and the ISB. It is hypothesized that:

1. There will be a greater frequency of acceptable needs reflected by the ISB than by the TAT.
2. There will be a greater frequency of unacceptable needs reflected by the TAT than by the ISB.

METHOD

Eighty male upper classmen college students were employed as Ss. These Ss were residents of two dormitories at Syracuse University, and had been assigned to these quarters, on a first come, first serve basis by the University officials. The Trait Rating Scale, the TAT, and the Rotter ISB were administered to 14 groups, the groups ranging in size from 3 to 11 Ss.

Three TAT cards, 4, 14, and 16, were selected for inclusion in this study. One view is that the reason is because they had been shown by Eron (1948) to produce themes reflecting the 4 experimental needs to roughly the same degree. Card 16 was selected because its bland nature would logically facilitate the highest form of ambiguity in the TAT media. As stated previously, the ISB series, when administered in its entirety, was considered as having no inherent set toward any single need.

The Trait Rating Scale was designed to provide a measure of the social acceptability of statements defined by Murray (1938; 1943), as indicative of the experimental needs. On this scale, the S had merely to indicate his willingness or reluctance to have a term used as "descriptive of him", by his social circle. Sample terms included ". . . is very loyal to his friends" (affiliation), and ". . . is

always ready for a fight", (aggression). An opaque projector was used to show the TAT cards on a 6 x 6 foot screen. The Ss sat behind and about the projector, all having a clear view of the screen, at distances of from 10 to 15 feet.

Prior to the administration of the TAT cards, the E gave the following directions derived from those recommended by Murray (1943).

This is a test of imagination, one form of intelligence. I am going to project some pictures onto the screen, one at a time, and your task will be to make as dramatic a story as you can for each. Use the blank sheets in the order in which they are numbered. Write only one story on each sheet. Write what has led up to the event, what is happening at the moment, and what the characters are feeling and thinking, and then give the outcome. There are three pictures and you have eight minutes for each. I will tell you when four and six minutes have passed and when you are to finish up, if you have not already done so. Here is the first picture.

These directions were given irrespective of whether the first card was 4 or 14. Card 16 was not given first at any time since it was felt that the Ss should first have some practice at the task, simplified by the employment of content material on the card. The possible orders of presentation were 4-14-16, 4-16-14, 14-4-16, and 14-16-4. No additional instructions preceded the other content card, but the blank card was always introduced as follows (Murray, 1943):

This is a blank card. Imagine some picture here and describe it to me in detail.

Although 120 Ss were to be employed in the original format, it be-

came impossible to obtain a sample of this size. Therefore, 60 Ss were given the ISB followed by the TAT; 15 Ss under each order of cards. Only 20 Ss received the TAT followed by the ISB, 15 under card orders 4-14-16, and 5 under card order 4-16-14.

The two raters were not informed of the purpose of the experiment. They knew only that they were to determine if one or more of four needs were reflected by a given statement according to the criteria given them. Rater A was a college graduate, having majored in psychology as an undergraduate, somewhat unfamiliar with projective techniques. Rater B was a second year graduate student in psychology, possessing considerably greater familiarity with the concept of projectives. The TAT stories were divided into individual sentences, and each sentence was judged plus or minus, depending on whether some expression of each of the four experimental needs was present. Because of this method, any single statement might contain as many as four or as few as zero needs. Each individual ISB statement was rated in the same fashion. The total number of times which a need appeared in the protocol was then divided by the number of scoreable statements in the protocol, yielding a percentage.

RESULTS

The testing of the basic hypotheses would not ordinarily require equated groups since in each instance it was the individual S's TAT test profile which was compared to his ISB profile. However, it may be that the

TABLE I—Experimental Design Providing for Control of Card Order and Test Presentation

Number Sequence of Cards	4, 14, 16	4, 16, 14	14, 4, 16	14, 16, 4
Cell Number	1	2	3	4
ISB followed by TAT				
Proposed N	15	15	15	15
Actual N	15	15	15	15
Cell Number	5	6	7	8
TAT followed by ISB				
Proposed N	15	15	15	15
Actual N	15	5	0	0

order of presentation of projective tests will affect the information elicited. Moreover, it is also possible that the sequence of TAT card order may have an effect on the protocols.

In order to evaluate these possible effects, the experimental design shown in Table I was employed.

This original design provided for comparisons between the effect of order of presentation of tests and the effects of card sequence. Multiple interactions could not be obtained because of the insufficiency of Ss in cells 6, 7, and 8. This lack occurred because the term ended and the students left for summer vacation prior to the conclusion of the test series.

As a prerequisite to statistical analysis, it was necessary to demonstrate that the groups were fundamentally identical prior to testing. Otherwise, differences arising might be a function of intrinsic group characteristics, extraneous to the test variables. Therefore, a chi square analysis was performed between the groups on the following variables: age, number of Ss having parents living and together, number of male siblings, and number of female siblings (McNemar, 1949; Siegel, 1956). Since cell 6 was comprised of only 5 Ss, it was combined with cell 5. Comparisons made throughout this study will be between cells 1, 2, 3, and 4, each of 15 Ss, and cell 5, of 20 Ss. Card order was confounded, but test sequence maintained.

The five experimental groups were not significantly different on any of the four criteria variables, and were therefore considered as being drawn

from the same population. Any difference which might be reflected between test sequence, then, is a function of this specific variable.

The next factor investigated was whether or not the needs which had been defined (a priori) as being socially acceptable or unacceptable were also seen in the same way by the sample population. Since there had been no variation in administration procedure for the trait rating scale, i.e., it was the first form given, irrespective of cell, all ratings were summed and *t*-tests performed between the means of each of the four needs.

The *t*-values shown in Table II demonstrate that the 80 Ss rated statements expressing needs Aggression and Succorance as significantly more socially undesirable than statements expressing needs Achievement and Affiliation. This comparison holds true in every instance where a socially acceptable need is compared against a socially unacceptable need. Neither of the acceptable needs, compared one to the other, was seen as being significantly different. That is, while need Aggression was seen as socially different from needs Achievement and Affiliation, it was not seen as being different from need Succorance. The same relationship holds for each of the other needs.

As an index of similarity between the judgments of the two raters, a percentage comparison was made for each need reflected by the ISB protocol and TAT cards 4, 14, 16. This was done for each sample cell. To the sum of the number of statements, upon which both raters agreed as con-

TABLE II—Means and *t*-Values of the Differences in Social Acceptability of Four Experimental Needs
(*N* = 80)

Need	\bar{X}	Affiliation <i>t</i>	Aggression <i>t</i>	Succorance <i>t</i>
Achievement	2.71	.08	3.57**	3.78**
Affiliation	2.57	...	3.66**	3.91**
Aggression	10.1700
Succorance	10.14

***p* < .01

taining a specific need, was added the sum of statements upon which they were in disagreement. This term was then divided into the total number of statements upon which the raters agreed, yielding a percentage.

Table III shows that the two raters agreed with one another to a fairly high degree in most instances. The percentage agreements compare favorably with those reported by previous investigators employing several types of rating scales to determine content on projective techniques. The percentage of rater agreement on the ISB ranges from 88 per cent on total need Achievement to 81 per cent on total need Succorance. When considering the combined reliability of the TAT cards, the agreement is somewhat lower, ranging from 81 per cent on total need Aggression to 68 per cent on total need Achievement.

In the light of these findings, the agreement between raters was sufficiently high to lend credence to the view that the criterion instrument, the definitions of the experimental needs, provided the same general scoring concepts to different raters.

The ISB percentage agreements tended to be closely grouped among the specific cells. Considering absolute magnitude, it might appear that need Achievement was better defined than need Affiliation, need Aggression or need Succorance, listed in descending order of rater agreement. Yet when the variance of cell percentages about any specific need is taken into account, it is apparent that the range of cell agreement percentages within any single need is greater than the range of the overall percentage of the four needs. Therefore, it appears that none of the experimental needs was

TABLE III—Percentage of Rater Agreements in Determining Four Experimental Needs for Two Projective Media as Occurring in Five Sample Cells

Need	Cell	N	ISB Percentage	TAT Card Percentages			
				4	14	16	Total
Achievement	1	15	91	75	61	76	65
Achievement	2	15	93	0	78	86	62
Achievement	3	15	90	60	70	55	62
Achievement	4	15	88	0	83	73	62
Achievement	5, 6	20	79	57	84	83	82
Total Achievement	Σ Pop.	80	88	32	74	78	68
Affiliation	1	15	80	86	33	83	78
Affiliation	2	15	89	73	69	100	80
Affiliation	3	15	84	64	40	83	75
Affiliation	4	15	85	60	100	80	76
Affiliation	5, 6	20	86	56	92	91	82
Total Affiliation	Σ Pop.	80	85	69	76	86	78
Aggression	1	15	80	86	83	87	85
Aggression	2	15	81	74	81	78	78
Aggression	3	15	88	88	74	85	82
Aggression	4	15	80	86	63	80	78
Aggression	5, 6	20	78	84	81	82	82
Total Aggression	Σ Pop.	80	82	84	78	75	81
Succorance	1	15	87	100	80	100	95
Succorance	2	15	93	26	73	100	52
Succorance	3	15	91	68	68	50	65
Succorance	4	15	71	83	88	17	63
Succorance	5, 6	20	70	78	78	50	76
Total Succorance	Σ Pop.	80	81	75	43	43	71

more precisely defined than the others, at least as employed in ISB scoring.

The over-all percentages of the combined TAT cards present a somewhat different picture. While no statistical comparisons were performed between these means, it is evident that the cell per cent of any single need cluster more closely to the sum per cent of means of that need, than do the sum per cent of means of the individual needs one to another. Therefore, some greater belief may be attached to the view that certain of the needs were defined more adequately than others as used in TAT scoring.

As a final note to this table, it will be noted that the variability of rater judgment is greater for the individual cells than for the sums of the three cards. It is felt that this phenomenon is not due to any intrinsic error in the scoring mode, the summing of the cards providing for an averaging-out of errors. Rather, it is believed that this effect is due to wide variability in the Ns of certain card-cells. To illustrate this, if the number of mentions of a need in a cell is as low as 5, one can easily see how a reliability of either 0 per cent or 100 per cent can be obtained. The Ns of scored statements for each cell for the ISB protocols and the TAT cards were examined and it was noted that the reliability coefficient was extreme only where the cell Ns were small.

A series of ten chi squares was calculated to determine if any of the subject cells differed significantly in productivity proportionate to the total number of responses, one from another, for any single experimental need. Stated another way, it was conceivable that any single sample might display a significantly greater amount of a need on either projective technique. Should this be the case, the cells have to be evaluated independently. A situation would have arisen which would necessitate the evaluation of five separate sets of data. On the other hand, if the cells were basically

similar in terms of proportional productivity there would be a statistical advantage because of the increased population N.

Since all chi square comparisons were two by two designs, a correction for continuity, recommended by Edwards (1957) was employed. Because of this, negative relationships were obtained in some instances. These are reported as .00, because of the one-tailed assumption inherent in the use of this technique.

Thirty-eight of the forty comparisons are non-significant. It appears, however, that cell 5 displays a proportionally greater amount of need Achievement than do either cells 1 or 2. This same factor does not hold true when cell 5 is compared to 3 and 4. This being the case, and since the cells had already been shown to be essentially similar in terms of the four variables shown in Table II, the cells were treated as if they were a part of the same population. All cells were summed when testing the basic experimental hypotheses.

It is known that chance fluctuations affect chi square, and this may well be an example of a type II error (Edwards, 1957). As a final note on this

TABLE IV—Chi Squares of Differences in Productivity Between Five Cells for Four Experimental Needs as Measured by the Incomplete Sentences Blank

Need	Cell	Cell			
		2	3	4	5
AFF	1	.18	.98	.00	.16
AFF	2		.00	1.09	.00
AFF	3			.89	.55
AFF	4				.60
ACH	1	.00	.57	.57	6.99*
ACH	2		.57	.57	6.69*
ACH	3			.00	3.37
ACH	4				3.37
SUC	1	.00	2.14	.00	1.37
SUC	2		.33	.00	1.49
SUC	3			2.14	.01
SUC	4				1.37
AGG	1	.54	.19	1.24	2.08
AGG	2		.00	.00	.47
AGG	3			.67	.43
AGG	4				.00

* $p < .05$

sequence, 2 of 10 comparisons may be expected to appear significant at the 5 per cent cut-off by the operation of chance factors alone.

The sum of the four chi square results was 14.96, far in excess of the critical values for the tests of proportionate productivity of needs on the TAT cards. A correction for continuity was again incorporated, and negative coefficients are similarly reported as being .00.

TABLE V. Chi Squares of Differences in Productivity Between Five Cells for Four Experimental Needs as Measured by the Thematic Apperception Test

Need	Cell	Cells			
		2	3	4	5
AFF	1	.57	2.17	1.25	.00
AFF	2		.75	.00	1.61
AFF	3			.00	1.13
AFF	4				.16
ACH	1	.39	.00	.39	.01
ACH	2		.00	.00	.01
ACH	3			.00	.00
ACH	4				.01
SUC	1	.14	2.13	.00	.39
SUC	2		.96	.72	.00
SUC	3			3.13	.96
SUC	4				.39
AGG	1	.00	.00	1.67	.00
AGG	2		.00	1.86	.99
AGG	3			1.86	.99
AGG	4				5.05*

* $p < .05$

Of the 40 comparisons made in this sequence, only one, the proportionate difference between cells 4 and 5 on need Aggression, is significant at the .05 level. As has been pointed out, the operations of chance factors alone would produce two chi square results in excess of the .05 level. It would appear that the expression of the four experimental needs is roughly of the same frequency for all groups. Returning to the performance of the Ss comprising cells on need Achievement, these scores are not different from those made by the Ss in the other cells. Greater confidence can be given to the method of summing the data by specific needs for the entire group. Presentation of the ISB prior to the TAT cards (or vice versa) does not

affect the results significantly in proportionate productivity.

The rationale for the statistical procedure involved in testing the basic hypotheses was as follows: if the two projective measures elicit basically the same sort of information, it would be expected that the percentage of need-mention by each measure would be roughly equivalent. Due to chance fluctuation, in fact, the two measures need not be equivalent. If the expression of needs is on one measure more than the other, according to the purpose, it was anticipated that a significantly greater number of Ss would demonstrate higher frequencies of need Achievement and Affiliation on the ISB than on the TAT. It was further anticipated that a significantly greater number of Ss would demonstrate higher frequencies of need Succorance and Aggression on the TAT than on the ISB.

While originally it was intended to compare any Ss ISB protocol against each of the TAT cards, the paucity of TAT statements, in some instances, tends to make such a comparison unstable. Therefore, the most accurate TAT need index was felt to be a pooled measure.

Prior to the final testing of the major hypotheses, each S's protocols were rated to determine if the frequency of any need was higher on the ISB or on the TAT cards. All disagreements in coding were discussed among the two raters and the author, and a final consensus reached. The results of the following table (Table VI) are reported on the basis of this final form.

All major hypotheses were given experimental validation. Affiliation and Achievement were reflected to a greater degree on the ISB than on the TAT. Conversely, Succorance and Aggression were shown to be present to a greater degree on the TAT than on the ISB. These findings were emphasized when the relation between ISB and TAT protocols was examined in terms of the sum acceptable and the sum unacceptable needs.

TABLE VI—Significance Value of Acceptable Needs Appearing Higher on the ISB than on the TAT and Unacceptable Needs Appearing Higher on the TAT than on the ISB for Eighty Ss

Need	N TAT Greater Than ISB	N ISB Greater Than TAT	Critical Ratio
Affiliation	31	49	2.00*
Achievement	20	60	4.47***
Succorance	49	31	2.00*
Aggression	67	13	6.06***
Sum acceptable	51	109	4.59***
Sum unacceptable	116	44	5.70***

* $p < .05$

*** $p < .001$

DISCUSSION

No direct comparisons can be made between this study and any others, since a review of the literature disclosed no directly related previous studies. There is a considerable difference between the basic drive type of experiment as Sanford's (1936a; 1936b), Murray's (1933), and Levine, Chein and Murphy's (1934), and this one employing social values. Restated, those studies may be best related to a Maslow schema (1954), whereas this would be most directly concerned with a framework akin to Murray's (1938). Murray's view (1938), that the TAT is a measure of needs which appear largely in thought alone in unmodified form, is not inconsistent with these findings. The study demonstrates that the magnitude of needs expressed is different on TAT stories than on at least one other projective measure. Yet, that this occurs because of a depth of personality factor is not a direct conclusion. Bellak (1959a) was likewise careful not to postulate exact origins of different need structures.

The experimental needs may have been gleaned from the same stratum of personality, yet because of the method, the TAT may be less susceptible to the censoring process seemingly operant when the ISB is used. Balken and Vander Veer (1942) may there-

fore be correct in contending that projectives measure on the same level.

Rosenzweig (1950) and Wyatt (1947) felt the TAT varies in depth from time to time. All that may be added is that this variation, if it actually occurs, is sufficient to change the need structure elicited in comparison to other projective measures.

Silverstein (1957) felt that needs will be expressed with greater strength when the anxiety attached to them is reduced. Less involvement in the stimulus material would likewise have the effect of lowering anxiety, since contact is less personal. Moreover, stimuli requiring less involvement reflect anxiety arousing needs in greater strength.

Both Combs (1947), who contrasted autobiographical material with TAT protocols and Gordon (1953), who contrasted dream content with TAT protocols, found that there exists a continuum of the expression of unacceptable aspects of personality. That is, autobiographical material contained more socially acceptable patterns than did TAT stories than did dreams. These results, contrasting two experimental measures, follow a parallel course and produce evidence that there is such a continuum.

Rotter's (1951) and Rhode's (1957) views are both reasonable if their meaning of the term "semi-projective" is eliciting material more consistent with the ego-ideal. Fully-projective tests (as the TAT) produce results less in line with the ideal self than do semi-projective tests (as the ISB).

Dorken's contention (1953) that projectives yield results with no necessary relationship may be questioned. Knowing the need pattern produced by one of the tests, a reasonable prediction may be made about the appearance of that pattern on the other.

Considering the theory and experimental design underlying this study, several aspects are worth noting. The question of exactly what constitutes the specific mechanism causing the difference between the two experimental measures is still unresolved.

For present purposes, an understanding of the interactions between the measures has a genuine heuristic value. While no gross predictions can ever be made without a basis in theory, there is practical value to pure qualitative prediction by knowledge of effects. On this basis, this study finds a niche, along the lines laid down by MacFarland *et al.* (1951)

An obvious failing in the study was the lack of subjects to fill cells 6, 7, and 8, as originally proposed. Previous work has hinted that the sequence of presentation of tests may have a significant effect on the results. No such difference was found, but this might well be a function of the Ns in the group given the TAT cards first, which was then followed by the ISB. Any future replication should attempt to provide an equated N sequence to demonstrate this difference. It is anticipated that this sort of design would point out that while differences may arise because of order of presentation, the same relationship of needs as here found will hold true.

Card sequence was altered more as a control measure, than as an attempt to point out differences from this manipulation. Testing effects of order was actually impossible, design-wise. It may be that the card order does have an effect on the patterns reflected by the TAT. Again, comparisons with Ns of 15 per cell are minimal for statistical purposes (Edwards, 1957; McNemar, 1949). Type I errors are very easily the resultant of such a procedure.

Fault may be found with the items concerned with detecting the social value of the experimental needs. These items were taken from Murray's definition (1943), but were quite unrefined by use in a pilot study. Moreover, there were only two behavior samples (i.e., descriptive statements) per need. Obviously a more extensive sampling is desirable. As partial compensation, an 11-point rating scale was employed. Upon examination, a truncated negatively skewed curve was

obtained for the social desirability attached to the acceptable statements. The social desirability of the unacceptable statements can best be described by a truncated positively skewed curve. The effect of this truncation was to mask the true differences. Therefore, it may be expected that the difference between the two sets of needs is even greater than the *t*-values shown in Table II.

Part of the difference in rater reliability may be a function of individual differences in sophistication in psychological procedures. Inductive evidence as to the effect of this difference in training was the consistently higher reliability in scoring between Rater B and the author, than between Rater A and Rater B or Rater A and the author.

As previously stated, statistical differences between productivity of selected needs between cells was felt to be an artifact of the general computational technique, and the assumptions inherent in the level of confidence concept. Only 3 of 80 comparisons were significant at the .05 level. By chance, 4 of 80 comparisons would be expected to be significant. Therefore, it was assumed that the Ss were drawn from the same population and the datum of the six groups was pooled for a common estimate.

There is a possibility that group 5 was actually different in expression of need structure from the remaining population. One alternative was selected, but it is realized that this position is by no means certain. No reason has been found to consider that the results of the study are significantly affected by this, since even if cell 5 is dropped from the final calculations, the difference between the expression of sum acceptable and sum unacceptable needs on these two measures is still in excess of the .001 level of significance.

SUMMARY

An experiment was conducted testing whether or not the magnitude of

certain needs reflected by a projective measure was a function of intrinsic characteristics of the test. Accordingly, the basis of this investigation was the conjecture that different projective media elicit basically the same sort of information. There is some question as to whether material elicited from different instruments is directly relatable even though they may all be subsumed under the heading "projective tests."

The subject population was eighty male upper classman students. Four experimental needs were selected; Aggression and Succorance being considered as socially undesirable for males, and Achievement and Affiliation being considered as socially desirable. Statements reflecting each of these needs were rated by the sample population as to their social acceptability, and desirable needs were shown to be significantly more acceptable than undesirable needs.

It was hypothesized that a test such as the TAT would reflect a greater amount of unacceptable needs and a lesser amount of acceptable needs than would a test such as the ISB. All comparisons were shown to be significant. Several possible explanations to account for this difference were considered. These included the possibility that the TAT provided a lesser amount of structure and/or that the subjects "identified" to a greater degree with the ISB items.

REFERENCES

- Balken, E. R., & Vander Veer, A. H. The clinical application of a test of imagination to neurotic children. *Amer. J. Orthopsychiat.*, 1942, 12, 68-80.
- Bellak, L. On the problems of the concept of projection. In L. E. Abt & L. Bellak (Eds.), *Projective psychology: clinical approaches to the total personality*. New York: Grove, 1959, 7-33.
- Bellak, L. Thematic apperception test in clinical use. In L. E. Abt & L. Bellak (Eds.), *Projective psychology: clinical approaches to the total personality*. New York: Grove, 1959, 185-229.
- Cattell, R. B. Principles of design in projective or misperceptive tests of personality. In H. A., & G. L. Anderson (Eds.) *Projective techniques*. Englewood Cliffs, N. J.: Prentice-Hall, 1951, 55-101.
- Combs, A. W. A comparative study of motivation as revealed in thematic apperception stories and autobiographies. *J. clin. Psychol.*, 1947, 3, 65-75.
- Dorken, H. J. Projective tests and the consistency of personality structure: a pilot study. *J. abnorm. soc. Psychol.*, 1953, 48, 525-531.
- Edwards, A. L. *Experimental design in psychological research*. New York: Rinehart, 1957.
- Eron, L. D. Frequencies of themes and identification in the stories of schizophrenic patients and non-hospitalized college students. *J. consult. Psychol.*, 1948, 12, 387-395.
- Gordon, H. L. A comparative study of dreams and responses to the Thematic Apperception Test; a need press analysis. *J. Pers.*, 1953, 22, 234-253.
- Holmopple, J. O., & Miale, F. *Sentence completion*. Springfield, Ill.: Charles C. Thomas, 1945.
- Kass, W. Projective techniques as research tools in studies of normal personality development. *J. proj. Tech.*, 1956, 20, 269-272.
- Levine, R., Chein, I., & Murphy, G. Relationship of the intensity of a need to the amount of perceptual distortion, a preliminary report. *J. Psychol.*, 1942, 13, 283-293.
- MacFarlane, J. W., & Tuddenham, R. D. Problems in the validation of projective techniques. In H. A. & G. L. Anderson (Eds.), *Projective techniques*. Englewood Cliffs, N. J.: 1951, 26-55.
- Maslow, A. H. *Motivation and Personality*. New York: Harper, 1954.
- McNemar, Q. *Psychological statistics*. New York: Wiley, 1949.
- Murray, H. A. The effect of fear upon estimates of the maliciousness of other personalities. *J. soc. Psychol.*, 1933, 4, 310-339.
- Murray, H. A. *Explorations in personality*. New York: Oxford University Press, 1938.
- Murray, H. A. *Thematic Apperception Test*. Cambridge: Harvard University Printing Office, 1943.
- Rhode, A. *Sentence completion method: its diagnostic and clinical application to mental disorders*. New York: Ronald Press, 1957.
- Rotter, J. B. Word association and sentence completion methods. In H. A. & G. L. Anderson (Eds.), *Projective techniques*. Englewood Cliffs, N. J.: Prentice-Hall, 1951, 279-311.
- Rotter, J. B., & Rafferty, J. E. *Manual: the Rotter Incomplete Sentences Blank-College Form*. New York: Psych. Corp., 1950.
- Rosenzweig, S. Some relationships between attitudes expressed toward the parents in a sentence completion test and case history data. *J. proj. Tech.*, 1950, 14, 188-193.
- Samuels, H. Validity of personality-trait ratings based on projective techniques. *Psy-*

- chol. Monogr.*, 1952, 66, No. 5 (Whole No. 337).
- Sanford, R. N. The effect of abstinence from food upon imaginal processes: a preliminary experiment. *J. Psychol.*, 1936, 2, 129-136.
- Sanford, R. W. The effect of abstinence from food upon imaginal processes: a further experiment. *J. Psychol.* 1936, 3, 145-159.
- Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- Silverstein, A. B. The expression of acceptable and unacceptable needs in thematic apperception. Unpublished doctoral dissertation, New York University, 1957.
- Windle, C. Psychological tests in psychopathological progress. *Psychol. Bull.*, 1952, 49, 451-482.
- Wyatt, F. G. Scoring and analyses of the Thematic Apperception Test. *J. Psychol.*, 1947, 24, 319-330.

Received December 22, 1961

An Experimental Study of the Effect of Stimulus Variation Upon Projection¹

EDITH WEISSKOPF-JOELSON AND HELEN C. FOSTER
Purdue University

It is the purpose of this study to throw some light on a question which may concern clinical psychologists attempting to develop new pictorial projective techniques or to select pictures from existing ones: what kind of pictures elicit the greatest amount of projection? Any consideration of this problem is not only of practical importance with regard to the selection of pictorial material, but also of theoretical importance with regard to the effect of stimulus variables on projection.

The subjects whose projections are being considered in this study are children of Kindergarten age. Two questions regarding stimulus variations were investigated:

(1) How do pictures showing animal figures compare with pictures showing human figures as to the amount of explicit verbal projection they tend to elicit? This question is one aspect of a more general investigation of the effect on projection of physical similarity between story-teller and pictorial central figure (Weisskopf & Dunlevy, 1952; Weisskopf-Joelson & Money, 1953). It is often assumed that animal pictures tend to be of higher diagnostic value for children than human pictures; the CAT is based on this assumption. Various researchers have attempted to test this hypothesis, but their investigations yielded contradictory results. In some studies the assumption was corroborated (Bills, 1950); but more often it was found that the stories told to human pictures are

equally productive (Biersdorf & Marcuse, 1954) or more productive (Armstrong 1954; Light 1954; Mainord & Marcuse 1954) of diagnostically useful material than the ones told to animal pictures. The present study differs in various ways from the previous ones. For example, our animal pictures differ from the human pictures only with regard to the figures; all other factors, such as action and background were kept constant. Bills and Light fail to introduce such controls. The children of the present study are younger than the ones used in any other of the cited studies. Weisskopf's (1950) Transcendence Index was used as a measure of productivity, as contrasted to the other studies, with the exception of Armstrong's. In contrast to Armstrong, we eliminated the carry-over effect from each human picture to the analogous animal picture and vice versa: each subject was presented with either the human version or the animal version of each picture, but never with both versions. Finally, a second factor, namely color, was varied in addition to the species of the figures, and the interaction between the two variations was considered.

(2) The second question regarding stimulus variation which we investigated was the following: How do colored pictures compare with achromatic ones as to the amount of projection they elicit? There has been little experimental effort to investigate this question. Thompson and Bachrach (1951) published one of the few studies on this topic. In general, their results fail to show any significant difference in productivity elicited by the two kinds of pictures. However, they indicated a tendency

¹The paper is based on a thesis presented to the Graduate School of Purdue University, Lafayette, Indiana, by the junior author in partial fulfillment of the requirements for an M.S. degree in psychology. The study was directed by the senior author.

of the colored pictures to be slightly superior in eliciting clinically valuable protocols.

The main hypotheses to be tested in this study are as follows:

(1) There is no statistically significant difference between the amount of projection elicited from Kindergarten children by pictures showing human figures and the amount of projection elicited by pictures showing animal figures.

(2) There is no statistically significant difference between the amount of projection elicited from Kindergarten children by colored pictures and the amount of projection elicited by achromatic pictures.

PROCEDURE

The materials consisted of modifications of each of four CAT pictures showing the following scenes: a rabbit lying in a bed, all alone in a large bedroom (Set 1); a lion sitting in a chair (Set 2); a kangaroo carrying groceries, with her two offspring (Set 3); and a dog spanking her offspring in the bathroom (Set 4).

Each card was used in four different versions; namely in a non-color version with animal figures (AN), a color version with animal figures (AC), a non-color version with human figures (HN), and a color version with human figures (HC).

When substituting human figures for animal figures and colors for black and white, we attempted to keep all other aspects of the pictures roughly constant. This control made it desirable to modify the CAT cards slightly, instead of using them in their original form. For example, the animals were shown with clothes, the kangaroo of CAT card #4 was shown carrying her offspring in her arm rather than in her pouch, and the mouse in Card #3 was omitted.

The subjects were 40 kindergarten children attending public schools at Lebanon, Indiana. Their ages ranged from five years, six months to seven

years, with a mean age of six years, two months.

Each child was presented with four pictures, i.e. one from each of the four sets, and from each of the four versions (AN, AC, HN, and HC). The selection was made in such a way that no child was presented with more than one picture from the same set, or of the same version. This selection was achieved by the use of a Greco-Latin square design, whereby one subject was assigned to each row of each square. Ten different squares were used, which were obtained by switching the columns, or the rows, of the original square according to random numbers.

The pictures were presented according to Bellak's suggestions regarding the administration of the CAT. The stories were recorded on tape.

RESULTS

Weisskopf's (1950) Transcendence Index was used as a measure of productivity.

A comparison of the mean Transcendence Index of all human pictures (HN and HC) with the mean Transcendence Index of all animal pictures (AN and AC) by a Student's *t*-test shows that the null hypothesis cannot be rejected at, or beyond the 5% level ($t = .995$).

Likewise, a comparison between the mean Transcendence Index of all achromatic pictures with the mean Transcendence Index of all chromatic pictures by a *t*-test yields no significant result at, or beyond, the 5% level ($t = .894$).

Table I gives the total of the Transcendence Indices for the four

TABLE I—Sum of Transcendence Indices for Each Set and Version

Form	Situation				Total
	1	2	3	4	
HC	92	78	72	71	313
HN	66	88	55	60	269
AC	71	99	62	87	319
AN	87	68	91	72	318
Total	316	333	280	290	1219

sets and the four versions. On the basis of an analysis of variance, the differences between the scores shown in this Table fail to reach significance at the 5% level. There is a trend for Set 3 (grocery scene) and Set 4 (bathroom scene) to elicit less projection than Set 1 (bedroom scene) and Set 2 (lion scene).

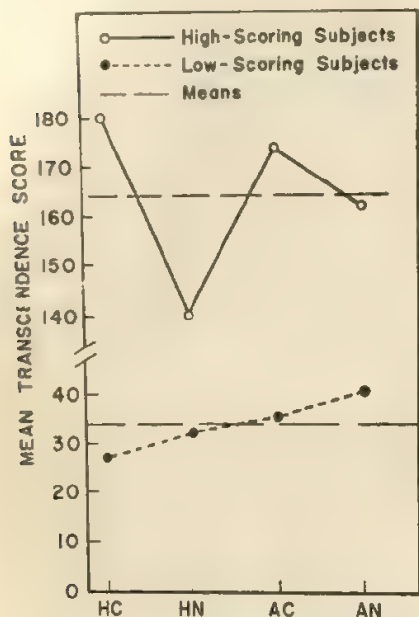


FIGURE 1—Sum of Transcendence Indices of High and Low Scoring Subjects for the Four Versions of the Pictures.

Figures 1 and 2 show the totals of the Transcendence Indices for the four situations and for the four versions based only on the protocols of the eleven subjects with the highest transcendence and the eleven subjects with the lowest transcendence. These extremes were compared in order to test the hypothesis that the effect of the human-animal variable, and of the color variable, upon projection might be different for subjects of high transcendence than for subjects of low transcendence, and that the failure to reject the null hypothesis might be due to a cancelling-out ef-

fect between the high scoring and the low-scoring subjects.

According to Figures 1 and 2, the low-scoring subjects tend to be more productive when responding to animal pictures than when responding to human pictures. Similarly, achromatic pictures tend to elicit slightly more productive responses than chromatic pictures with low-scoring subjects. The high-scoring subjects tend to show greater productivity when responding to chromatic pictures than to achromatic ones. None of the above differences is significant at, or beyond, the 5% level on the basis of t-tests.

Finally, it should be briefly noted that Set 4, the bathroom scene, elicits the highest productivity among the low-scorers, and the lowest productivity among the high scorers.

DISCUSSION

The data on the performance of

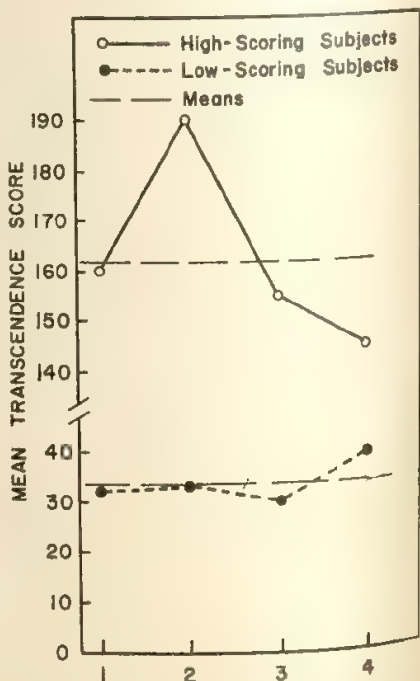


FIGURE 2—Sum of Transcendence Indices for High and Low Scoring Subjects for the Four Sets of Pictures.

the Kindergarten children used in this experiment support both our hypotheses. They do not support the impression of many clinicians that children respond more productively to animal pictures than to human pictures. This conclusion is in agreement with studies such as Biersdorf & Marcuse (1954). Similarly, this study did not show that chromatic cards differ from achromatic cards as to their effect on productivity.

The trend for Set 3 (grocery scene) and Set 4 (bathroom scene) to elicit less projection than Set 1 (bedroom scene) and Set 2 (lion scene) may be related to the fact that the children being studied are within the oedipal age range and can, therefore, be expected to be less preoccupied with oral and anal themes (grocery scene and bathroom scene) than with themes revolving around sleeping alone (bedroom scene) and relating to a father figure (lion scene). This interpretation relates the findings shown in Table I to Weisskopf's (1950) findings that subjects of college age are significantly more productive when responding to pictures about parent-child relationships and about hetero-sexual relationships among contemporaries than to pictures about other topics. Both observations lend some support to the hypothesis that people are most productive when responding to pictures about the major preoccupations of their age group.

The trends shown in Figures 2 and 3 appear plausible in terms of theoretical considerations. High-scoring subjects are more free to develop and express fantasies than low scoring ones. The latter may be more guarded and repressive with respect to their emotions. Thus, the diagnostic significance which clinicians ascribe to responses to chromatic stimuli would lead to the prediction that high-scorers should show more transcendence on chromatic cards than on non-chromatic ones and that the reverse would be true for low-scorers. Simi-

larly, low-scorers, i.e. subjects who are relatively inhibited when revealing information about themselves, may find it easier to do so under the pretense that they are speaking about animals rather than about humans.

The fact that the bathroom scene elicits the highest productivity among the low-scorers, and the lowest productivity among the high scorers, may be connected with the expectation that children with obsessive compulsive tendencies should be predominantly among the low-scorers rather than among the high-scorers; compulsives tend to stick to the factual evidence in the picture. At the same time compulsives tend to have anal preoccupations. Thus, the above data are in line with the previously stated assumption that pictures about major preoccupations tend to elicit relatively productive protocols.

SUMMARY

It is the purpose of this study to help determine whether projections elicited by pictures from the CAT were modified by substituting human figures for animal figures, and colors for black and white. The pictures showing human figures were compared with the ones showing animal figures regarding the amount of projection elicited. Likewise, the colored pictures were compared with the black and white ones. The results indicate that neither the animal-human variable, nor the color variable affect the productivity measured by Transcendence Indices. There is a tendency for subjects with low transcendence to tell their most productive stories to animal pictures and to achromatic pictures, and for subjects with high transcendence to tell their most productive stories to chromatic pictures.

REFERENCES

- Armstrong, M. A. Children's responses to animal and human figures in thematic pictures. *J. consult. Psychol.*, 1954, 18, 67-70.
Biersdorf, Kathryn, & Marcuse, F. L. Responses of children to human and animal

- pictures. *J. proj. Tech.*, 1954, 17, 455-459.
- Bills, R. E. Animal pictures for obtaining children's projections. *J. clin. Psychol.*, 1950, 6, 291-293.
- Light, B. H. A comparative study of a series of TAT and CAT cards. *J. clin. Psychol.*, 1954, 10, 179-181.
- Mainord, Florence R., & Marcuse, F. L. Responses of disturbed children to human and animal pictures. *J. proj. Tech.*, 1954, 18, 475-477.
- Thompson, C. & Bachrach, A. The use of color in the TAT. *J. proj. Tech.*, 1951, 15, 173-178.
- Weisskopf, Edith. A transcendence index as a proposed measure in the TAT. *J. Psychol.*, 1950, 29, 379-390.
- Weisskopf, Edith & Dunlevy, G. Bodily similarity between subject and a central figure in the TAT as an influence on projection. *J. abn. soc. Psychol.*, 1952, 47, 441-445.
- Weisskopf-Joelson, Edith & Money, L., Jr. Facial similarity between subject and central figure in the TAT as an influence on projection. *J. abn. soc. Psychol.*, 1953, 48, 341-344.

Received October 24, 1961

BOOK REVIEWS

Bennett, Edward. *Personality Assessment and Diagnosis*. New York: Ronald Press, 1961

Edward Bennett began his intellectual strivings at the Massachusetts Institute of Technology, where he received the B.S. in Electrical Engineering. He received the M.S.E.E. at Purdue in 1949. Moving to psychology he took the M.S. and Ph.D. in Psychology in 1951. During this strenuous program he instructed in Electrical Engineering.

From 1951 to 1959 he taught psychology at Tufts University. During part of this period (1957-1959) he was Associate Director of the Bio-Mechanics Laboratory at Tufts. In 1959 he published *The Search for Emotional Security*. Since 1959 he has been associated with the Human Factors Department, MITRE Corporation.

Doctor Bennett's current work introduces an "original technique for eliciting evidence of subjective feelings by means of multiple forced-choice judgment" which he labels The Polydiagnostic Method (PM). 18 chapters suffice to present this thesis. The first three chapters discuss human responses (interoceptive and exteroceptive) to environmental stimuli within a broadly psychoanalytic framework modified to a greater or lesser extent by elements of learning theory, field theory, and communication theory. Concepts are summarized in a series of graphic representations of emotional action and reaction which seems analogous to Newtonian physical constructs.

Part II, including Chapter Four through Six present the method of assessment. Entitled "Mapping Subjective Feelings", the procedure involves forcing the selection of three of 15 adjectives which *least describe* self feelings. The subject is then forced to select the three adjectives which *again least describe* the self from the remaining 12 words. This procedure is repeated until a set of three adjectives remains. The process continues through six sets of 15 items, followed by four sets in which the procedure is modified so that adjectives which *most describe* the self are chosen. Three sets of 15 adjectives are devoted to *what a person should be to be happy and satisfied*. Three sets of items are presented for selection in terms of *least describes people in general*.

The remaining two sets of the 20 item-lists in the battery require the subject to choose terms that *most describe people in general*. The subject's responses are summarized and represented graphically in order to "read the map".

Chapters Seven through 11 present three cases using the Polydiagnostic Method, with varying amounts of supportive case history and psychological test data. Chapters 12 and 13 discuss research design and methodology in general, with specific application to the PM technique. Chapters 15 through 18 give theoretical and clinical implications of research with this technique. Four Appendices include supplementary data on clinical cases presented, foreign equivalents of the polydiagnostic terms (French, German, Spanish, and Flemish), and a considerable body of normative data from the standardization sample.

The frequently equivocal nature of this work is perhaps best characterized by Doctor Bennett's statement in the preface:

"So many have contributed ideas to this book that to acknowledge them all would turn a general presentation into a bibliographical survey. Therefore, I have adopted the way of the physical and biological sciences; I have assumed that major ideas are now in the nature of general scientific information and may thus be discussed without mentioning the names of the contributors. In this way a personal omission is intended more as a compliment to a man's ideas than as a disservice to his person."

Data are handled in a variety of intensities, from means and variances for each item in the PM test to a perfunctory "As the pieces come together, some of these interpretations are in obvious disagreement; and they can be discarded". Hypotheses and conclusions are presented in an enthusiastic and decisive manner, while cautions are discernable, but soft-spoken. There are too many sweeping generalities, and unsupported enthusiasms to consider this a significant research contribution. It stands rather as an enthusiastic preliminary report. The word "enthusiastic", to the reviewer, seems appropriate in describing the author's style. Strikingly naïve and repetitious if intended for fully trained psychologists it would, on the

other hand, seem hopelessly complex and unrelated as a graduate text. However, Doctor Bennett writes with clarity and energy. One would look forward to his report of cross-validation with the PM technique, or perhaps to the application of his pleasant style to some of the better-known but muddled communicated clinical techniques.

THEODORE H. BLAU
Tampa, Florida

Honkavaara, Sylvia, *The psychology of expression*, Cambridge University Press, 1961; British Journal of Psychology, Monograph XXXII, pp. 96. \$4.50

The author attempts to show "how the developmental facts appear in the field of perception as a sequence of different modes or dimensions." This is an experimental study aimed at showing "how the dimensions in perception develop and increase in number as the human individual himself develops." How successful her attempt is will most likely be judged by the reader's analysis of her experimental designs and the statistical analysis growing out of these designs.

Four dimensions of perception are described, detailed and experimentally investigated. These are: 1. A dynamic-affective attitude which is conditioned by the state of the perceiver rather than the object itself. 2. A matter-of-fact relationship with the environment which is the ability to perceive objects as they exist in reality. 3. A physiognomic mode which is concerned with the perception of emotional expressions in others. 4. An intersensory property which is called sensitivity and best represented by the recognition of style in the outside environment. These are said to appear in developmental sequence. This particular monograph deals primarily with the third dimension—physiognomic expression.

The author is to be highly complimented on her chapter outlining the historical development of the problem. Seldom do we see such painstaking efforts directed at a research of past literature. It provides the reader with an excellent review of the literature concerning arguments over whether expressions are innate or learned. She paints in bold relief the generally held American position of their being learned as contrasted with the European position of their being innate. Her research is then directed to upholding the European viewpoint.

Honkavaara deals with the method of approach using observers to judge the various expressions seen in photographs. She is directly concerned with the question as to whether our ability to judge emotional expressions is innate or if we learn it through practice and social habits. She makes the point that in order for emotional expression recognition to be innate, it does not mean the recognition cannot realize itself in a sequential form. It need not appear all at once to be called innate.

The series of experiments she offers in evidence have one very refreshing aspect—they are not conducted on a college population. She uses children from three different cultures which gives her study added meaning. After having cited experimental results which, according to the author, showed that the recognition of expressions is a "relatively late development in the child", she proceeds to offer additional experimental findings demonstrating the different developmental levels prior to the recognition of expression.

The thinking of Honkavaara is similar to that of Piaget in many respects, and those who disagree with Piaget will most likely disagree with Honkavaara. I believe anyone interested in the field of child psychology would do well to read through this monograph not only for its provocative thought but for its ingenious experimental designs.

HAL M. WELLS
Elmira College

Sherman, M. H. (ed.) *A Rorschach Reader*. New York: International Universities Press, 1960. \$7.50.

The editor and the publisher of *A Rorschach Reader* have performed a valuable service for both beginning students and experienced practitioners in making available this collection of twenty-eight (28) significant papers on the most-widely used of the projective techniques. While there are available several excellent monographs on the administration, scoring, and interpretation of the Rorschach test, there are all too few single volume collections of those important papers which have appeared in the general literature, and Sherman's book is indeed welcome.

The papers contained in this 440-page volume have with one exception been published elsewhere. They are grouped under five categories with a well-written foreword statement by Ruth L. Munroe which establishes the frame of reference for the book. Sherman

himself has written a brief introduction to each paper which does a good job of presenting its essence and highlighting its importance.

The opening section of eight papers is devoted to Rorschach findings in psychopathology. Included are such classics as Evelyn Hooker's contribution on male homosexuality in the Rorschach which appeared originally in this journal in 1958. Hooker's paper is one of the basic studies in Rorschach test diagnosis of male homosexuality. When first published, Seymour Fisher's paper on "Rorschach Patterns in Conversion Hysteria" was a welcome addition to the literature since there had been relatively little in print dealing with this prominent syndrome. There are also included several papers dealing with schizophrenia and the Rorschach, including Bertram Forer's, "The Latency of Latent Schizophrenia," Florence Neally's, "Rorschach Sequence Analysis in a Case of Paranoid Schizophrenia," Margaret Mercer's, "Diagnostic Testing in Two Cases of Schizophrenic Depression," and Maria Rickers-Ovsiankina's, "Longitudinal Approach to Schizophrenia Through the Rorschach Method." This latter group of papers forms an interesting overview of method and findings in research via the Rorschach test into the question of schizophrenia. George Kisker's paper on "The Rorschach Analysis of Psychotics Subjected to Neurosurgical Interruption of the Thalamocortical Projections" includes a well thought out review of early Rorschach work in determining signs of organic brain impairment. As Kisker points out, organic pathology in the Rorschach can be inferred only against the background of the total psychological picture. He warns against the all too common error of anticipating the diagnosis on the basis of several so-called pathognomonic signs. The final paper in this section is one by David Shapiro on special problems of testing borderline psychotics, which constitutes one of the most widely discussed problems in both testing and psychotherapy today. His case illustrations are particularly valuable.

The second section of this anthology includes five papers dealing with Rorschach findings and vocational choice. Anne Roe's significant work in this area is represented by two papers on "Painting and Personality" and "Analysis of Group Rorschachs of Psychologists and Anthropologists." The former study in particular which was done in 1946 helped to launch an entirely new field of research employing projective methods. The study of creativity and talent is one which is presently receiving much attention, and

due acknowledgement must be made of the role which studies such as these have played in initiating this current interest. Walter Myden's paper on "An Interpretation and Evaluation of Certain Personality Characteristics Involved in Creative Production" has not been published to date elsewhere, but it is a valuable addition to the literature. He has referred to problems in sex-role identity in the creative personality and has related this point to Kris' concept of regression in the service of the ego. Myden's study varies somewhat with Roe's findings but the samples are not comparable, and there is a variation in the method of analysis of the data. Joseph Levi contributes a study on "Rorschach Patterns Predicting Success or Failure in Rehabilitation of the Physically Handicapped," which contains many good and original thoughts. He reflects in his writing a deep appreciation of the relationship between the Rorschach test and psychoanalytic theory. This paper, incidentally, is one of the few in the collection which was not originally published in either the *Journal of Projective Techniques* or its predecessor, the *Rorschach Research Exchange*. Another paper which is especially timely at the present is on the "Use of the Rorschach in the Diagnosis of Teacher Effectiveness" by Percival Symonds and Stephanie Dudek. Of importance for students of the Rorschach is that the authors show due appreciation for the fact that a certain trait or function may be reflected by many different aspects of the test.

Section three includes a trio of papers on Rorschach findings among other cultures. Jules Henry is represented by his pioneering study on "Rorschach Techniques in Primitive Cultures." This research using Pilaga Indian subjects is basic and is one with which every psychologist should be familiar. Philip Cook's article on the application of the Rorschach test to a Samoan group was published some twenty years ago and is a contemporary of Henry's study, as is A. I. Hallowell's "Acculturation Processes and Personality Changes As Indicated by the Rorschach Technique." This series of papers constitutes an excellent introduction to the use of the Rorschach and other projective techniques in the study of other societies.

Next follows a series of articles on the experimental use of the Rorschach. This is somewhat of a misnomer, inasmuch as nearly all of the other papers in this volume are also based upon experimental work. "Rorschach Content Analysis: An Experimental Investigation," by the British psychologists, Joseph Sandler and Bryan Ackner, offers an

interesting approach to the problem posed in their title through the use of factor analytic methodology. It seems to this reviewer that too many of the American studies on content analysis are either anecdotal or based upon a limited number of subjects, and the Sandler-Ackner paper offers an intriguing approach to the question, even though one may not agree completely with their presentation. Ruth Munroe has contributed one of the exceptional articles in the literature in a study relating Rorschach test findings to different modes or levels of intellectual functioning. Her findings are stimulating and can be related to and compared with the previously mentioned studies by Roe and Myden.

The final and fifth section of the book is concerned with the interpretation of test results. Zygmunt Piotrowski has a basic paper on movement responses as indicators of changes in personality which was originally published in 1937 but which has much relevance today. A. Furrer adds a paper on "The Meaning of M in the Rorschach," which is a translation from *Imago* in 1925. Yet, this article has an excellent contemporary tone. Furrer explores the nature of the creative impulse and the relationship between imagery and motor functioning. His article may be seen as an early contribution to ego psychology in psychoanalytic theory. There is

E. G. Schachtel's paper on "Color and Affect," which is another widely recognized classic which suggests that the relationship between Rorschach determinants and the formal elements in dreams could well be a profitable area of study. Fred Brown contributes a paper dealing with psychodynamic factors in Rorschach content analysis. Roy Schafer has an interesting paper on "Transference in the Patient's Reaction to the Tester" which has relevance not only for psycho-diagnostic work but for psychotherapy as well.

There is an extensive bibliography of eight pages which covers all references used in the individual papers. There is also a valuable subject index, as well as a name index listing all persons mentioned in every paper and all references. For the research student, there is a large number of good leads for further studies, and this is one of the valuable contributions of Sherman's book, in the opinion of this reviewer.

Over-all, this large book of many individual papers offers valuable aids to anyone working in the area of clinical psychology in general and psychological personality testing in particular. The choice of papers obviously has been a careful one.

MELVIN A. GRAVITZ, PH.D.
2025 Eye Street, N.W.
Washington 6, D.C.

ANNOUNCEMENTS

The Board of Trustees of the University of Louisville has approved a doctoral training program in clinical and experimental psychology. Graduate work beyond the master's degree will be initiated in September, 1963. Inquiries regarding the program should be sent to the Department of Psychology and Social Anthropology, University of Louisville, Louisville, Kentucky. Assistantships and stipends to \$3,000.00 are available in 1962-63.

BOOKS FOR REVIEW

The following books are available for review. We ask that those who offer to write reviews take seriously their responsibility to complete the review within three months. Write to Bertram R. Forer, Ph.D., Suite 307, 8833 Sunset Blvd., Los Angeles 69, Calif.

Ames, Louise B. & E.g. Frances L. Mosaic patterns of American children.

Bricklin, Barry; Piotrowski, Zygmunt, & Wagner, Edwin E. The Hand Test.

Wagner, Edwin E. The Hand Test: Manual and materials.

Klopfer, Bruno & Davidson, Helen. The Rorschach techniques: An introductory manual.

Muller, W. H. & Enskat, A. Graphologische Diagnostik.

Witkin, W. A.; Dyk, R. B.; Faterson, H. F.; Goodenough, D. R., & Karp, S. A. Psychological differentiation.

WORKSHOPS

The New York Society for Projective Techniques held four workshops in February and March.

1. The Bender-Gestalt as a projective technique. Fred Brown.

2. The Psychoanalytic interpretation of the Rorschach. Erwin Singer.

3. The clinical application of projective techniques with children, Helen Anderson, Tatania Juzak, H. Michal-Smith, Rose Wolfson.

4. Use of play therapy, puppetry and plastic media as a projective device and in diagnostic formulation. Adolf G. Woltmann.

CONVENTION

SOCIETY FOR PROJECTIVE TECHNIQUES, Inc.

Saint Louis, Missouri

1. Wednesday, Aug. 29th
1 PM-9 PM
Executive Board Meeting
Society for Projective Techniques
Pine Room, Ambassador Hotel
2. Thursday, Aug. 30th
4:00-4:50
Business Meeting
Colonial Room,
Chase Hotel
3. 6:00-6:50
Social Hour, Rudibunda Room,
Chase Hotel
4. 7:00-9:00 PM
Dinner, Colonial Room,
Chase Hotel
5. 9:00-10:30 PM
Presidential Address,
Edwin S. Shneidman,
"Projections on a Triptych: A Hagiology for Our Times"
Colonial Room,
Chase Hotel
6. Symposium: University Responsibility in Clinical Techniques Training.
Friday, August 31, Old English Room, Ambassador Hotel, 10:00-11:50 A.M.
Gordon F. Berner, Kenneth Little, William E. Henry, Donald Ford, Hugh B. Urban, Henry P. David, Michael H. P. Finn (with Div. 12)
7. Symposium: Using Projective Techniques in Understanding and overcoming Resistance to Change.
Friday, August 31, Old English Room, Ambassador Hotel, 1:00-2:50 P.M.
Harry V. McNeill, Pauline G. Vorhaus, Walter G. Klopfer, (with Div. 12).
Bertram Forer, Jules Holzberg.

Sale on Back Volumes

The Journal of Projective Techniques has a limited number of copies of each of the volumes published from 1947 to the present. For a short time we are willing to dispose of them at cost to individuals and below cost to libraries and other institutions (U. S. and foreign).

ORDER BLANK

Society for Projective Techniques, Inc.
210 E. Wilson Ave.
Glendale 6, California

Enclosed find \$_____ for the following order:

Bill me _____ for the following order:

Volumes	Year	Institution \$4.00	Individual \$5.00
11 & 12	1947-48	_____	_____
13	1949	_____	_____
14	1950	_____	_____
15	1951	_____	_____
16	1952	_____	_____
17	1953	_____	_____
18	1954	_____	_____
19	1955	_____	_____
20	1956	_____	_____
		\$4.50	
21	1957	_____	_____
22	1958	_____	_____
23	1959	_____	_____
24	1960	_____	_____
25	1961	_____	_____
Current Volume 26 at \$8.00			
(Foreign \$8.50)			
Monograph No. 1 at \$1.50			
"Projective Techniques and Psychotherapy" by William Lundin			
Monograph No. 2 at \$2.50			
"Manual for Make-A-Picture Story (MAPS) Method"			
by Edwin Shneidman			
Case of El reprints (complete) from			
Vol. 25 at \$2.50			

Name or Institution: _____

Street _____

City _____ Zone _____ State or Country _____

JOURNAL OF PROJECTIVE TECHNIQUES

VOL. 26

DECEMBER, 1962

No. 4

. . .

EDITOR

BRUNO KLOPPER

EXECUTIVE EDITOR

BERTRAM R. FORER

ASSOCIATE EXECUTIVE EDITORS

JEANNE M. REITZELL
EDWIN S. SHNEIDMAN

STAFF EDITORS

EVELYN CRUMPTON
FLORENCE DIAMOND
NORMAN L. FARBEROW
HERMAN FEIFEL
FRED J. GOLDSTEIN
JAMES HOLT

WALTER G. KLOPPER
LEONARD KOROT
JEROME REISEL
NICHOLAS ROSE
STANLEY F. SCHNEIDER
JOHN WEIR

IRLA LEE ZIMMERMAN

EDITORIAL:

A rose by any other name

The Board of Trustees of the Society for Projective Techniques has for several years been much preoccupied with the nature of the Society's role in clinical psychology. Originally the Rorschach Research Exchange, the Journal of Projective Techniques was created as a means of gathering and communicating experimental and clinical lore about the Rorschach method of personality diagnosis, a decidedly new kind of technique and methodology for personality description and diagnosis.

With the growing acceptance of the Rorschach came a proliferation of theory and techniques of assessment and diagnosis. The Rorschach Research Exchange added Projective Techniques and in 1950 became the Journal of Projective Techniques. Clearly and much to the disquiet of many psychologists projective techniques have been not only accepted but overbought in clinical psychology particularly, though not exclusively. The original purpose of the Society and its journal has been almost too successfully achieved. Now we are asking ourselves whether we should now fade gracefully from the scene. Has the Society a function in clinical psychology? Should we search for new goals?

Organizations tend to become self-perpetuated often beyond their period of usefulness. To invent goals in itself seems a denial of real purpose. The members of the Board did not have to search far for goals. We found them in our own thinking about our clinical work particularly in diagnostic testing. We are agreed that projective techniques are useful and that they are only part of a broader activity, the clinical psychological appraisal, assessment, and/or diagnosis of human

beings. We believe that projective techniques should not stand alone nor in opposition to other methods of clinical investigation. In fact, to employ one or more projective methods alone in diagnostic appraisal is to delimit arbitrarily and inappropriately our view of the person. Projective techniques need to be separated from the methodology of clinical inference and diagnosis which has developed around them.

In recent years we have broadened the scope of the journal's content by paying more attention to the contributions of the authors who have submitted manuscripts to us. We have been trying to formulate what has occurred in an effort to further deliberately the trends created by the Zeitgeist.

We intend now to accept manuscripts which employ non-projective as well as projective techniques: psychometric tests, behavior observations, interview methods and any other methods and techniques that are used in dynamic fashion to add to our sophistication in psychological appraisal. But our emphasis will be less on the techniques per se than on the process of utilizing them for the understanding of persons or groups.

While the journal is defined by what it published rather than by its name, it seems appropriate that the name be changed to reflect more adequately what it is about. We are deliberating about a possible future name.

We want now to invite our readers and contributors to help us to define and implement our current role in clinical psychology.

BERTAM R. FORER
Executive Editor

Projections on a Triptych; or a Hagiology for Our Time¹

EDWIN S. SHNEIDMAN

While roaming aimlessly through a large and wonderful museum recently, I came upon—in a dim cul-de-sac apart from the main gallery—a particularly fascinating triptych. Its three panels, each with the representation of a sainted figure, instantly caught my eye. Being in a reverent mood, it captured my interest, and I, having that day an active spirit that was susceptible to random stimulation of this kind, found my imagination excited by its mysteries. The presentations in this short paper are, in part, a result of the ruminations initiated by that triptych.

In that I aspire to deal with some abstractions in this paper, I shall, for purposes of contrast, begin by being specifically concrete, that is, by showing you that very same triptych, which, to use a euphemism, I "borrowed" from the museum at a moment when my favorite guard, who happens to wear a hearing aid, had momentarily turned his back. Permit me to add that I fully intend to return it at the earliest possible opportunity. Here it is. (See Figure 1).

Looking at the triptych, you will notice that the left panel depicts St. Carl Gustav, patron of printing. He is the special saint of ink and printers' tools. His domains of benign influence include workers with ordinary and special inks, bottled ink, spilt ink, and, of course, ink blots. He is concerned with printers' types of all sorts, including daguerreotypes, linotypes, and archetypes. He is the protector of the special graphic arts, and

thus one can refer to him as the saint of the idiographic. We shall have more to say about him.

The panel on the right is, as I hope you can see, one of St. Icarus. He is the saint of the air, of breezes and of gales. He is especially endeared to balloonists and those of us who purchase insurance policies in airline terminals. His interests include all things gaseous and aerial: hurricanes, cyclones, vacuums, winds, flat tires, pneumatic drills—all of pneumatics, and he is therefore often called the saint of the pneumathetic, which, as you well know, is usually mispronounced as nomothetic.

The middle, and perhaps most important, saint is St. Herman. He was an early divine who is noted for his preaching among the mountains and near the lakes, and is known to have died an untimely death after delivering his ten tablets—each remarkable for its ambiguity—to an eager world. Those of you who are versed in the history of our Society know the story well. But I have not indicated what he is saint of. In that this aspect of his character is crucial to the development of my theme, I shall hold this secret for awhile until a more auspicious moment for the revelation of his complete identity.

We have, at this point, partially identified our complement of these extraordinary personages. In this fruitful triptych we have the saint of the idiographic, the saint of the nomothetic, and the saint of aspects as yet unknown. Now our task is to see in what directions our own projections on this brief hagiology can lead us.

Let us begin with the saint of the nomothetic. Notice, in the figure on the right, the elegance, exactitude and precision of his garb. There is an

¹Presidential address delivered at the annual meeting of the Society for Projective Techniques, St. Louis, Missouri, August 31, 1962. This paper was written while the author was a U.S. Public Health Service Special Research Fellow (1961-1962), at the Department of Social Relations, Harvard University, studying with Dr. Henry A. Murray.

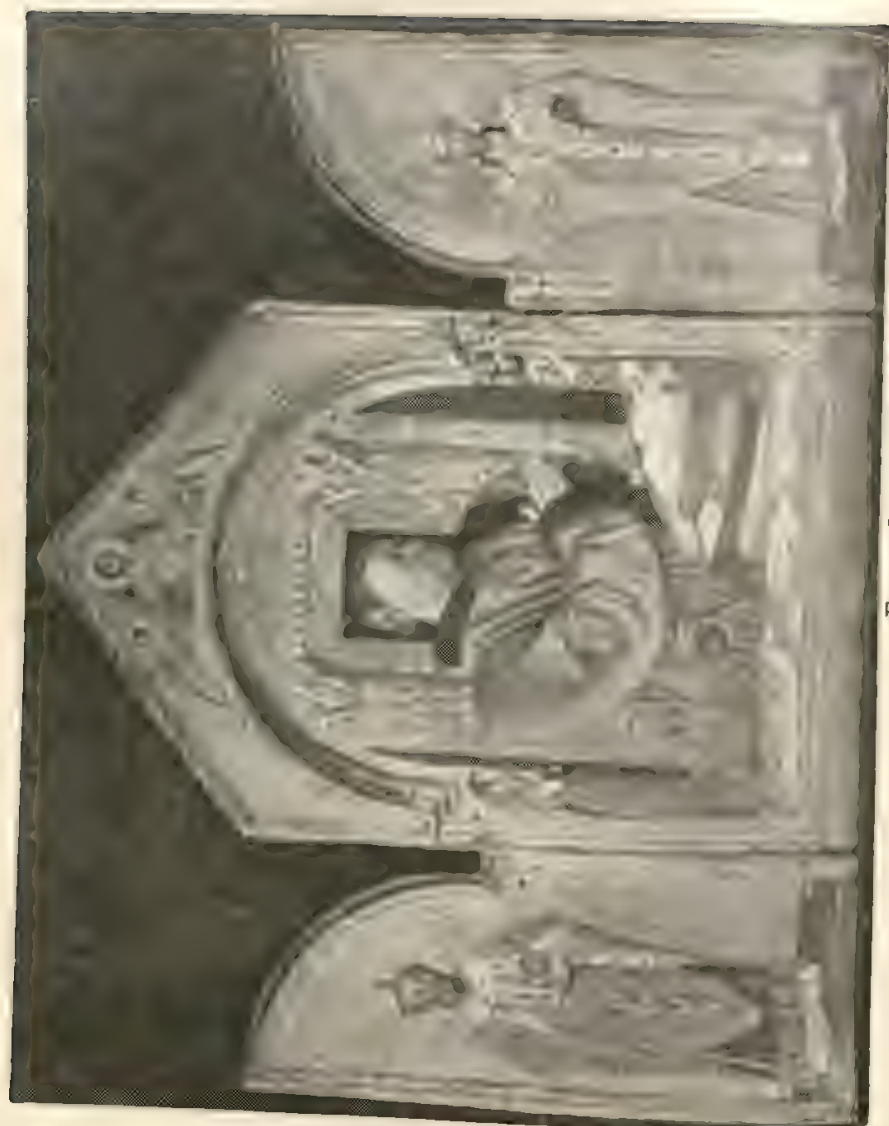


FIGURE 1

apocryphal story that the book that he is holding contains tables of random numbers; another story holds that it is much more organized and is filled with mean heights and weights, average IQ's, median M's and W's and so forth, for every major country in the world, divided, of course, for sex and pathological diagnostic category. His sermons are a joy to hear. Maybe not a joy exactly, but they are remarkably distinct and clear. He never says a sentence except that he takes pains to say it exactly the same way on another occasion. He can be depended upon—for that. It is said that he speaks an elegant Coptic but there are very few of us who can apply what he says to the pressing problems of our everyday lives. Some of us would, to confess it outright, trade some of his precision and reverence—and it might be added that his reverence seems to be for his precision—for a little more that is practical and relevant to the activities of ordinary (and extraordinary) human living. His published summarization, in Latin, of the average proportions of the fifty cathedrals in our country was inexplicably impressive, but in our secret hearts we would have appreciated more his discussing with us — in vulgate—how to put the roof on our own church. His concern with all of God's creatures is remarkable, but, lamentably, almost to the exclusion of his interest in man. For some reasons (wondrous to speculate on), he is especially devoted to the dog and the rat, although he has also been known to keep pigeons. For him, the proper study of mankind is math. But, withal, we are happy to say that we have many good friends among those who are devout members of his parish.

On the other hand, our idiographic saint has, as one might suspect, something of a different personality. It is said that on more than one occasion he has refused to take up a collection because he did not believe in aggregates. In his long walks through the country-side with his neophytes he

will often divine the meanings of cloud formations in such a way as to border on the miraculous, but he does not (or cannot) tell exactly how it is done, and it has been noted (no doubt only by disaffected detractors) that he has, more than once, made rather different pronouncements to rather similar situations. Inasmuch as all of his interpretations seem so sensible and meaningful, this unforeseen plethora of riches introduces the embarrassing problem as to which of his quite different prescriptions to believe. It might be added as an aside, that a hurried consultation with our nomothetical saint did not resolve the dilemma, for he was able to supply us only with information about cloud formations in general, how they varied as to season, altitude, and potentiality for precipitation. Our idiographic saint sticks pretty closely to humans, and seems especially interested in the confessional. He appears to believe that every situation is absolutely unique, and, even further, that some situations are more unique than others. He has great self-confidence and undiminished zeal in relation to his own beliefs. Once, for example, he was asked if he believed in baptism. "Believe in it?" he is reported to have replied, "Why Good Lord, man, I've seen it!" He is very diligent by nature, and has been known to devote himself for days and weeks to the study of the written biography of a single individual, feeling that if he understood this one person he could thereby know something of great importance about the particularities and universalities in every man. It can be added that even our nomothetic saint is fond of him but thinks that he is hopelessly (but romantically) muddle-headed, and tells him so. In light of this last fact, it is somewhat difficult to understand why, when one of the brothers of the nomothetic saint suffered from phobias and nightmares, he sent him, for counsel, to his idiographic colleague. The ways of the world—and of saints—are marvelous

to behold and oftentimes impossible to fathom.

At this point, the big question in my mind is whether or not we must limit our choice between these two side panels. There seem to be some whose words would rather imply that we must. But you will note that when these esteemed persons ask us to choose between these two—the clinical and the actuarial—we are able to do just that (and to ameehliorate the situation), not by picking one or the other as they would seem to imply that we must (and indeed they even suggest which one we ought to choose), but by selecting a position *in* between the two side panels. This type of choice permits us to avoid a totally unnecessary dichotomy and allows us to concentrate on the *quality* of the conceptual nature of our normative or individually-oriented enterprises.

As a small addendum, one can note too that the dichotomizers, it turns out after further thought, may not have been correct even in their initial premise regarding this polarity, for we see that Michael Polanyi, in his *Study of Man*, reminds us that Windelband (who, in 1894, thought up these notions in the first place in order to distinguish between the apparently unique character of historical events and the seemingly repetitive character of events in the natural sciences) did "not say that the field of reality can be divided into the subjects of nomothetic and idiographic knowledge," but rather said something quite different, namely that these two are forms of knowledge, and are, so to speak, simply "two logically distinct parts of all knowledge." Logically, there may be other logical positions.

Further, we can note that other conceptualizations of these different emphases of approach have been formulated. That most eminent Professor of Philosophy, the late Morris Cohen, in his review of Dewey's *Essays in Experimental Logic*, in-

sightfully combined philosophers and scientists and then described two main species of either, as follows: "Among philosophers as among scientists we may roughly distinguish two types that may be called the mathematical and naturalistic. The mathematicians excel in grasping some fruitful idea and elaborating it with such a perfection or finality of form that humanity is compelled, through sheer admiration, to strain the facts to make them fit these perfect forms. The naturalists, on the other hand, are more eager to observe the actual facts in their naked natural state. They love accuracy more than elegance. That philosophers can write with their eye on the object of their observation rather than on the symmetry of their final system, careful readers of Aristotle and Kant know full well. Nevertheless, the prevailing temper among philosophers has been the mathematical one; for all the great men of science whose achievements have stirred the human imagination, from the days of Euclid and Archimedes down to Copernicus, Galileo, and Newton, have been mathematicians. It was only in the middle of the nineteenth century, and most notably in the case of Darwin, that men of the naturalistic type succeeded in impressing humanity with results of the first magnitude."

The general point that we wish to make at this juncture is this: rather than to argue the relative merits of the idiographic or nomothetic approach (or to decide the merits of an approach in terms of whether it is essentially idiographic or nomothetic), it would make much more sense to inquire about the "goodness" of the conceptualizations. If poor concepts abound, neither idiography nor nomothetics will be fruitful: if rich concepts are present, it may not then be critical whether the enterprise is primarily one or the other, in that the quality of the concepts may superordinate the label.

A brief personal digression may be

useful here. As some of you know, I have, for the past year, been at Harvard University enjoying the great pleasures of talking to and listening to a number of their illustrious professors in such varied fields and disciplines as social relations, psychology, philosophy, human development, public health, public administration, psychiatry, etc. In casting my mind back over the memories of this year's experience, I have taken delight in thinking *ad hominem*, characterizing each of these eminent teachers in terms of a famous historical personage. Thus, in the contemporary Harvard scene, in the departments I have mentioned, I have found one professor who espouses a Cartesian point of view, one who is a present-day Plutarch, one who attempts to emulate Ebbinghaus; there is a Leibnitz, a Mahan, a deTocqueville, a Breughel, a Linnaeus, and so on. Now, in line with the *ad hominem* approach, one is not precluded from making personal comparisons among these worthies. But how? It seems to me that the most relevant dimension is one of *timeliness*. Did not Aristotle himself say that it is the mark of an educated man to know the degree of precision (referring to both observation and report) that is most appropriate to each state in the development of each realm of knowledge? And, if one may quote a Yale man in this context, do we not have the dictum of F. S. C. Northrop "that to talk about scientific method apart from specification of the specific stage of inquiry for a given type of problem is . . . meaningless"? Who then, *in our field*, are for our time and who are anachronistic? Who are most appropriate, useful, meaningful in terms of the paramount scientific needs of our time? One must first ask, what are these needs? Is it not true, as that great modern-day Harvard Linnaeus—who believes that psychology is the scientific study of *man* and that "the marked deficiency in our supply of ordered knowledge about

compound types of normal human beings is that psychology in its anxious, importunate determination to become an accredited experimental science, skipped the basic naturalistic stage of careful observation, description and classification which has constituted the first chapter in the development of all other sciences that deal with different varieties of entities"—has said that it is most appropriate for psychologists at our present stage of development to be "man-watching naturalists"?

For those of us who feel more comfortable with the identification and status that accrues to a party label, who have been debating between conservative idiography and radical nomothetics, the banner of liberal personology—with its specific humanistic orientation—has a great deal to recommend it. The personologist selects from both extremes. He is not content to know some few pre-selected facts about many people, or many facts about only one person, but rather he aspires to investigate intensively several people—paranormal, normal, supernormal—seeking what is common to all, present in some, and unique to each, all the while keeping his inner eye open for new groupings, new families or idiosyncracies, new human complexes. The personologist is bold enough to modify the traditional courtroom oath: he swears to investigate the individual, all the individual, but stops short of limiting himself to nothing but the individual, for he includes the individual in his natural habitat and in his pacific and stressful dyadic relationships. Further, the personologist uses—or feels free to select among—all the available psychological tests that meet his purposes of immediate observation and subsequent classification. He feels equally free to invent and construct new devices, instruments, techniques and procedures as the situation and the terrain demand. He is concerned as much with the creative analysis of conversations within the social hour

as of consultations within the therapy hour. He searches out opportunities to engage in bold theoretical reflections as well as in rigorously sophisticated researches; he tries to direct himself away from the trivially clear and the clearly trivial. These flexibilities of approach, taken together with his goals (of studying human subjects in meaningful ways) may permit him to ford swollen streams of consciousness, cross chasms between dichotomous bluffs, avoid jungles of mazes, traverse deserts of scientific aridity, and get over mountains of information—on the long march to the promised lands. He is, as you can see, an adventurer as well as a pioneer.

I wish to reiterate my belief concerning our optimal primary orientation. We are not atomic physicists, not Boolean algebratists, not experimental biologists, not particulate physiologists, not symbolic logicians—but rather we are, or ought to be, man-watching naturalists.

Our prayer is for more modern Linnaeuses and for more Linnaean attitudes in all of us. It seems that we, in sheer sibling-emulation of sister sciences have shown unnecessary celerity in trying to close the doors of our incompletely constructed taxonomic barns. The great difference between us and some of the physical and other biological sciences is that they have their cows and sheep and horses inside the barn, counted and classified, whereas our barn is both somewhat empty and the clusterings of the inhabitants are definitely confusing. Along with the nonsense of locking the barn door after the horse is stolen is the equal foolishness of freezing the shape of the incubator before the unhatched chickens are counted.

The Board of Trustees of the Society has, as part of its on-going responsibilities, given serious thought to the useful extension of the role and scope of the Society. One of the suggestions seriously entertained was that of redesignating our organization as

the Society for Projective and Assessment Techniques. The idea has considerable appeal, but I am still bedeviled by the feeling that I do not satisfactorily know how our subjects project, or how we, as psychologists, assess. I might almost be tempted, if this were the beginning of my tenure rather than its conclusion, to propose instead that we consider calling our group the Society for Paradigmatic Techniques. Such a designation might be more accurate and more challenging, for one of the ways that we might face our reality is by recognizing that much of our reality is abstract.

What does it mean to believe in paradigms? For one thing it means that we abandon our belief in many of the melancholy Santa Clauses of contemporary psychiatry and clinical psychology. These not-so-happy mythical figures masquerade under such names as alcoholism, delinquency, schizophrenia, suicide, homosexuality, and so on. The Santa Claus belief is that all adults are either homosexual or heterosexual—that all cases of schizophrenia—whatever that is—are hebephrenic, catatonic, paranoid, or simple (sometimes complicated with such modifiers as sub, latent, pre, incipient, borderline, chronic, acute, pseudo, etc.). We all know that schizophrenia has been variously divided into two types, six types, four types; that schizophrenia has been thought of as metabolic (Kraepelin); paleological (Arieti); ego deficiency (Alexander); emergence of the archaic (Jung); habit disorganization (Meyer); projection of repressed homosexual desires (Freud); etc. And even this brief recitation omits the more recent notions of schizophrenogenic mothers, autonomic balance, psychotic plasma, reticular formation, and the drug of the week. Clearly, we cannot treat and cure until we can achieve understanding; clearly, the need is to set aside conceptualizations of little usefulness and to strive for powerful concepts and for viable taxonomies.

One additional illustration (of the near-uselessness of some of our current keystones) may suffice. This example has to do with suicide. In regard to "suicidal" aspects of personality, what we do currently is to measure our Rorschach, TAT, Blacky, MAPS test, Szondi, House-Tree-Person and anamnestic protocols against the criteria of "non-suicidal" versus "suicidal"—the latter sometimes being further subdivided among commit, attempt, and threaten,—often overlooking whether our subject did, is, will, won't; often gainsaying all the complexities of the individual's belief systems about death, hereafter, rebirth; often neglecting his particularized ways of reasoning and his idiosyncratic logical styles; often ignoring his special need systems and his individual constellation of press; often forgetting that he is engaged in dyadic relationships of great variety and nuance; often closing our eyes to a multiplicity of possible unique psychodynamic ferments and ego organization; often side-stepping the issues of ego-passivity and ego-activity. Such delinquencies of omission seem to be part and parcel of our most honored current enterprises. For my part, if I know anything about suicidal phenomena it is that the present-day crude classification (or assemblage of gross categories) is essentially confusing, if not downright obfuscatory. This is one reason why some of my own efforts in this field in the past year have been directed toward the attempt to develop new classifications and fresh models to embrace many behaviors now loosely labeled as suicidal. Very briefly, I have been concerned specifically with such concepts as cessation, termination, interruption, and continuation, particularly the variety of "orientations towards cessation" which can be conceptualized in terms of the individual's intention—but this is not the appropriate setting for an explication of those thoughts. However, it is appropriate to remark that classifications of ob-

served and elicited phenomena under these (and other relevant) rubrics would be a necessary beginning to meaningful understanding—and ultimately to preventing, controlling, predicting, treating—these now only dimly understood self-destructive and other-effecting behaviors.

I hope that I have made the point—really two points are implied: one, that in our field there is no *ding*, only data; and two, that, further, we know no pure-color data, only hypothesis-saturated data. We have no choice—regardless of our personal intolerance for ambiguity and our concomitant need for structure, reliability, precision, elegance, pseudo-certainty—but to subsume our professional lives under the purview of hypotheses, analogies, paradigms, and constructs.

It would appear, at first thought, that what I am saying is that individuals in clinical psychology and psychiatry have to deal primarily with the methodological problems attendant to *construct* validity. But, even on this very issue, let us see what our second thoughts contain.

We all know the proposal regarding types of validity so lucidly enunciated by the APA Committee on Test Standards in 1952, a classification which included predictive validity, status validity, content validity, and congruent or construct validity. (There is a new committee presently at work on this formulation, addressing itself to a remodeling of this decade-old formulation).

Our re-reflections about construct validity might lead us to reaffirm its crucial role in our enterprise, but as we examine the concept itself we might well conclude that it contains two major shortcomings. The first is that the phrase "construct validity" might turn out to be an interesting contradiction in terms. It is, if you see what I mean, like saying "unrestorable truth," or "debatable fact," or "transient verity," or "unsubstantiated belief," or "unreal reality." That is to say, to conceptualize in terms of con-

struct validity is tantamount to stating that one insists that the phenomena under investigation correlate high with ideas that one somehow trusts and with taxonomies, usually implicit and un verbalized, whose ordering of the universe into apparent phenomena (implied by the taxonomy) one has unconsciously or uncritically accepted as either optimal or real.

A second major shortcoming latently contained in the concept of construct validity is the possible general confusion between validity and other types of verifiability; that is, the danger of assuming that an explication of types of validity exhausts the topic of "truth." The propositions within contemporary psychology appear to be amenable to division among a relatively few major different degrees of confirmation, conceptual buttressing and/or empirical verification. Thus, within psychology, there appears to be a certain range, containing perhaps no more than three or four more-or-less similar degrees of precision of prediction and exactness of repeatability (ranging from unprovable to plausible) compared with the larger number of precision-uncertainty positions within the totality of sciences and disciplines. Recognizing this, we should, in our activities and pursuits, be continuously committed to correcting our speculations, insofar as is possible, against facts and data. Further, we may have to extend our thoughts about psychological truth not only to include the notions of degrees of certainty, precision, and completeness, but also to encompass the concepts of espousal, proof, and maximum verifiability, along with the ideas of validity. One implication of this is that a classification of types of validity (which might well include something akin to construct validity) would be only one sub-section in a large array of several ways of cognizing our findings and taking our data to our bosom.

It needs to be added that the outmost reaches of verifiability which any data can possess may be set by the theoretical notions with which they are identified. In this sense, it may be that the maximum possible "validity" of the Rorschach technique, for example, as most of us who use it currently conceptualize it, may be somewhat less (but certainly not one whit greater) than the present veridicality status of the psychoanalytic personality concepts (like super-ego control, latent homosexuality, regression in the service of the ego, etc) which currently give the Rorschach technique its theoretical vitality and its intellectual permissiveness.

If "validity" is a concept that we wish to preserve, it may be most useful at this time to employ a straightforward description of each particular study in terms of a rather simple common-sense classification of (a) validity from *whom* (describing the subjects, the examiners and the judges in terms of the variables deemed to have major relevance); (b) validity for *when* (in terms of future variables or prediction, past variables or postdiction, or more-or-less present variables or what I call paridiction); (c) validity for *which* (in terms of either factual variables or inferred variables); (d) validity for *what* (in terms of the stimuli which were employed and the continua or dimensions along which it is believed that those stimuli are deployed); and (e) *how* validity was established (in terms of the statistical and other manipulative procedures employed). This brief spelling-out of the process for each study would increase initial understanding and facilitate replication by making explicit the concepts (in the stimuli, in the personnel, in the situation, etc.). It would underscore one main point of this presentation: that we recognize the abstract and essentially paradigmatic quality of our constructs.

Retrospecting over our brief trail, it would appear that we have attempt-

ed to build markers at four separate points. They are these:

1. The current view of the psychodiagnostic field as idiographic or nomothetic is unnecessary dichotomous, limiting our vision to the two side panels and focussing our attention away from the center ring of contemporary thought.

2. Whether idiographic, nomothetic, or whatever, most of our psychological realities are concepts—constructs, metaphors, analogies, paradigms.

3. Some of us have greater need for certainty and structure than others. One way of seeking to achieve certainty is by way of validity, and that type of validity which apparently has most relevance for the psychodiagnostic enterprise is construct validity. But the concept of construct validity may itself imply a contradiction in terms in the sense that one cannot have non-reality (constructs) and certainty (validity) as part of a single encompassment. The concept of construct validity may reflect too much of the nomothetist's need for preciseness at the expense of relevance.

4. It is believed to be axiomatic that before one can meaningfully and efficiently treat, predict, help, one has to understand, and, further, that the heart of understanding depends upon meaningful taxonomy or classification. Our great current need is for multiple viable classifications. In relation to this orientation, the position of the personologist has much to recommend it. This position holds that psycholo-

gists should be "man-watching naturalists," finding, describing, and systematizing the multitudinous individual and social behaviors of their fellow human creatures.

But we have yet to talk about St. Herman. What kind of a saint is he? What is he saint of? Here was a mystery indeed. Permit me to share with you the process whereby this puzzlement was solved. By examining the painting carefully, I discovered that there were, on the podium in the middle panel, two small silver coin-like objects. Accordingly, I took the trip-tych to a Professor of Classical Numismatics—an elderly Englishman whose hobby was the coining of phrases—who, to my great delight, identified each of those two objects as a denarius, that is, a Roman coin worth exactly ten cents. In light of the fact that there were two of them, he suggested that the figure in the panel be called the Saint of the Pair o' Dimes! And so he is: the saint of analogies, allusions, allegories, abstractions, models, patterns, conceptualizations; the saint of if as-if then but-then! He is, I propose, the patron saint of our Society.

As a last thought, I wish to state that the Society for Projective Techniques should make haste slowly by manifesting autonomy, initiative, integrity, maturity, at its own rate; and that this rate should be set at a pace which is consistent with its own nature, its own ego, its own needs, its own development, and its own identity.

Saints preserve us!

Psychosexual Development of Stutterers

FRANCES M. CARP
Trinity University

The Blacky Pictures (Blum, 1949, 1950) have been used to assess psychoanalytic theory in regard to sex differences (Blum, 1949), paranoid schizophrenics (Aronson, 1953) and ulcer patients (Blum and Kaufman, 1952; Blum and Hunt, 1952; Bernstein and Chase, 1955; Berger, 1959). Here predictions from psychoanalytic theory are checked against Blacky Pictures scores for stutterers.

METHOD

Hypotheses. The *Psychoanalytic Theory of Neurosis* by Fenichel (1945) was selected as the statement of psychoanalytic theory for this study. According to Fenichel "the symptom of stuttering reveals more readily than other conversion symptoms that it is the result of conflicts between antagonistic tendencies; the patient shows that he wishes to show something and yet does not wish to. Since he consciously intends to speak, he must have some unconscious reason for not wanting to speak . . . For (stutterers), the function of speech regularly has an anal sadistic significance . . . one may speak, in stuttering, of a displacement upward of the functions of the anal sphincters." (pp. 311, 312). Fenichel continues "that anal-sadistic wishes play the most prominent role in stuttering does not mean that other eroticisms or component impulses take no part in it . . . three component impulses usually play a characteristic part in the symptom of stuttering: The Phallic, the Oral, and the Exhibitionistic." (p. 313). For clarity specific predictions regarding Blacky Picture scores are presented below with results relevant to them.

Subjects. Blacky protocols were obtained from 20 stutterers and 20 non-stutterers attending a midwestern col-

lege. Each stutterer was matched with a non-stuttering student of the same sex, age, family constellation, and classification on the Ohio College Aptitude Test (OCAT). There were 15 men and five women in each group. Age ranged from 19 to 39 years with a median age of 22. On the OCAT, one pair was in the D classification, eight in the C, five in the B, and one in the A. Eight pairs were "only" children; the largest number of siblings for any pair was five.

Data Collection and Processing. Standard group testing procedures were used (Blum, 1950). Protocols were coded, those of stutterers mixed with those of non-stutterers, and scored according to Blum's system as modified by Aronson (1953). Scores were matched, each stutterer's to his control's, and differences were tested by means of chi-square for correlated scores (McNemar, 1955) using an arbitrary cut of Very Strong and Strong vs. Weak and Absent. Chi-square was not computed when any expected frequency (E) was less than five; Yates' correction for continuity was used when any E was less than 10. (McNemar, p. 231).

For two of the scoring categories, Preference and Related Comments, one of the E's was usually zero and only once as much as five, and no chi-squares were computed. Preference scores are determined by selections of the single "best" and "worst" of the 11 cartoons. Cartoons X and XI are preferred so commonly these responses are not scored. Similarly, a response scorable on one variable of the test was made spontaneously to a different cartoon so infrequently that Related Comment scores could not be treated statistically.

The scoring categories remaining are not directly comparable. The

Overall Dimensional Score (ODS), based on all subscores, has the advantage in regard to reliability. Nevertheless subscores which could be treated statistically, Spontaneous Story (SS) and Inquiry (I), were included because of their suggested relationships to different levels of psychological functioning. According to Blum "Interpretation of spontaneous stories is dynamically oriented, with emphasis upon latent rather than manifest content" (1950, p. 6), while "the multiple choice items of the Inquiry, being the most highly structured part of the technique, tend to evoke responses at or close to the conscious 'level'" (1950, p. 8).

RESULTS

Primary Impulse: Anal Sadism. "Psychoanalysis of stutterers reveals the anal-sadistic universe of wishes as the basis of the symptoms. For them, the function of speech regularly has an anal-sadistic significance. Speaking means, first the utterance of obscene, especially anal, words and, second, an aggressive act directed against the listener . . . The expulsion and retention of words means the expulsion and retention of feces, and actually the retention of words, just as previously the retention of feces, may be either a reassurance against possible loss or a pleasureable autoerotic activity." (Fenichel, 1945, p. 312).

In regard to anality, the core factor in the psychoanalytic explanation of stuttering, Blacky Pictures are scored for Retentiveness and Expulsiveness. However the two sets of scores are obtained from responses to a single cartoon, and scoring high on one anal scale reduces the possibility of scoring high on the other. It seemed advisable because of this scoring artifact to deal with only one. Since Retentiveness appeared to be the decisive factor, the prediction was for stutterers to score higher than non-stutterers on it.

Results are presented in Table I. In Anal Retentiveness, stutterers did not

differ significantly from non-stutterers in any scoring category. Test findings did not agree with the hypothesis which was stated for the primary impulse.

On Anal Expulsiveness stutterers scored lower (this score is 5.00 vs. significant at .05 level) than non-stutterers on ODS. Greater Retentiveness in stutterers might show up, because of lack of independence of these categories in the Blacky, as an insignificant increase on the Retentive scale and a larger decrease on the Expulsive. Perhaps Anal Expulsiveness is the relevant scoring dimension, and the prediction should have been for depressed scores due to severe depression.

Secondary Components. "Three component parts usually play a characteristic part in the symptom of stuttering: the phallic, the oral, and the exhibitionistic." (Fenichel, 1945, p. 313).

Phallic. "The function of speech is frequently connected with the genital function, particularly with the male genital function. To speak means to be potent; inability to speak means castration. Boys frequently reveal that their eagerness to talk well has developed as a substitute for phallic competition . . . Girls with a similar ambition have the unconscious wish to function genitally like men." (Fenichel, 1945, pp. 313-314). Higher scores in Castration Anxiety (males) and Penis Envy (females) for stutterers were predicted.

Stutterers obtained higher scores than non-stutterers on all three of the

TABLE I—Chi-square Comparisons of Stutterers with Non-Stutterers

	SS	I	ODS
<i>Primary Impulse</i>			
Anal Retentive	2.75	1.59	3.65
<i>Secondary Components</i>			
Phallic	6.72*	6.18*	7.24*
Oral Erotic	4.46*	.40	1.67
Oral Sadistic	4.99*	1.74	6.25*

df = 1 in all tests

*significant at the .05 level

Blacky scoring categories, in complete agreement with the prediction.

Oral Eroticism. "In less severe cases of stammering, communication with objects has been sexualized, entirely or in specific associations, and therefore speaking has become disturbed. In more severe cases of stuttering the function of communication has been given up entirely; the organs of speech are again intended to be used autoerotically." (Fenichel, 1945, p. 315). Higher Oral Erotic scores were predicted for stutterers.

A significant difference was found only in Spontaneous Stories, which Blum suggests tap "deeper" levels of personality functioning. Evasion was important in determining this difference: significantly more stutterers gave unusually short answers (one simple sentence) or gave short relevant answers and went on to discuss or describe something extraneous ("Mother feeding Blacky. Mother appears to be wearing lipstick and has a bow about her neck. She also has flashy-looking eyelashes. The slide is cracked.")

Oral Sadism. "The stutterer may not only unconsciously attempt to kill by means of his words; at the level now being discussed his symptom also expresses the tendency to kill his words, as representing introjected objects." (Fenichel, 1945, p. 314). Higher Oral Sadistic scores were predicted for stutterers.

Blacky Pictures responses confirmed the hypothesis on Overall Dimensional Scores and Spontaneous Story. The Inquiry, perhaps indicative of relatively surface level functioning, showed no difference.

Exhibitionism. There is no directly relevant Blacky scale. However, "the unconscious aims of the objectionable exhibitionism are not only direct erogenous ones; as in perverse exhibitionism, the reaction of the audience is needed for a reassurance against castration anxiety or as a satisfaction of some narcissistic need; and the means by which this reaction is de-

manded may unconsciously be very sadistic. Because of the unacceptable sadistic way in which the reassurance is sought, the castration fear, against which reassurance was needed is still further intensified . . ." (4, p. 316). The participation of exhibitionistic tendencies should reinforce Castration Anxiety and Penis Envy scores. As had already been noted, scores for these Blacky Pictures variables were significantly higher for stutterers in all three categories.

DISCUSSION

Test results agreed in no scoring category with the prediction for the primary impulse. Blacky Pictures do not provide scores directly relevant to one secondary component (exhibitionism), but for the three scorable on the Pictures (phallic, oral erotic and oral sadistic), test results were consistent with predictions in six of nine comparisons at the .05 level of confidence. In every case, Spontaneous Story scores supported the psychoanalytic prediction for secondary components.

Predicted differences between stutterers and non-stutterers clearly exceeded chance. In regard to secondary components, results are satisfactorily consistent with psychoanalytic theory. Lack of agreement between prediction and scores on the primary component cannot be interpreted clearly. The theory may be wrong; the prediction may be incorrectly drawn from theory; or agreement between theory and responses may be obscured by a scoring artifact in the Blacky Pictures. Independence of the Anal Retentive and Anal Expulsive scales of the Blacky Pictures would help to clarify such situations.

SUMMARY

Predictions concerning personality characteristics of stutterers were derived from psychoanalytic theory and tested by means of Blacky Pictures scores. Results were consistent with predictions in regard to secondary

components but not in regard to primary impulse. Changes in the Blacky Pictures scoring system are needed to clarify the meaning of this disagreement of scores with prediction.

REFERENCES

- Aronson, M. L. "A study of the Freudian theory of paranoia by means of the Blacky Pictures." *J. proj. Tech.*, 1953, 17, 3-19.
- Berger, L. "Cross validation of 'primary' and 'reactive' personality patterns with non-ulcer surgical patients." *J. proj. Tech.*, 1959, 23, 8-11.
- Bernstein, L., & Chase, P. H. "The discriminative ability of Blacky Pictures with ulcer patients." *J. consult. Psychol.*, 1955, 19, 377-380.
- Blum, G. S. "A study of the psychoanalytic theory of psychosexual development." *Genet. Psychol. Monogr.*, 1949, 39, 3-99.
- Blum, G. S. *The Blacky Pictures*. New York: Psych. Corp., 1950.
- Blum, G. S., & Hunt, H. F. "The validity of Blacky Pictures." *Psychol. Bull.*, 1952, 49, 238-250.
- Blum, G. S., & Kaufman, J. B. "Two patterns of personality dynamics in male peptic ulcer patients as suggested by responses to Blacky Pictures." *J. clin. Psychol.*, 1952, 8, 273-278.
- Fenichel, O. *The Psychoanalytic Theory of Neurosis*. New York: W. W. Norton and Company, 1945.
- McNemar, Q. *Psychological Statistics*. New York: John Wiley and Sons, 1955, (2nd ed.).

Received April 17, 1962

A Comparison of the Human Figure Drawings of Psychoneurotics, Character Disturbances, Normals, and Subjects Experiencing Experimentally-Induced Fear

JOHN E. EXNER, JR.
Bowling Green State University

Sundberg's (1961) report that the Machover Draw-A-Person Test (1949) is the second most commonly-used instrument in clinical work in the United States clearly indicates that a large number of clinicians find the device useful. At the same time, much of literature concerning the diagnostic usefulness of the technique seems to yield negative rather than positive findings. Swensen's (1957) review of literature pertaining to the technique led him to the general conclusions that many of Machover's basic hypotheses concerning the interpretation of human figure drawings have not been supported nor is there any strong experimental evidence to support the clinician's general estimate that the technique is a worthwhile clinical instrument. Swensen cites many studies which offer conflicting findings concerning several of the basic aspects of drawings often used for interpretation; such as line quality, shading, placement on the page, implied movement, and the like.

The present study was originally designed to give major emphasis to the phenomena of anxiety as it may or may not manifest itself in human figure drawings. As the study progressed, however, many additional elements worth reporting were noted so that the total data collected, although gathered from small samples, does seem reasonably representative of four basic groups: clinically-diagnosed psychoneurotics, clinically-diagnosed character disturbances, subjects manifesting experimentally-induced fear, and subjects clinically examined and found to be essentially normal.

Two hypotheses were formulated: First, that anxiety associated with a

psychogenic disturbance, such as that found in the psychoneurotic and character disorder, would manifest itself in the line qualities of human figure drawings in such a way as to be significantly different from the drawings of essentially normal subjects; Second, that subjects experiencing experimentally-induced fear would manifest this fear in the line qualities of human figure drawings in a manner similar to psychoneurotic and character disturbance subjects.

PROCEDURE

Data were collected from four groups of 20 subjects each, each group containing 14 females and 6 males between the ages of 18 and 20. The subjects in the first group, Group N, were all self-referred clients to a university clinic. In the intake interview they complained of chronic and intense anxiety and frequent and intense depressions. Using a test battery which included the Rorschach and Thematic Apperception Test, all were diagnosed as psychoneurotic and fell into three basic sub-categories of obsessive-compulsive, phobic, or depressive. All were eventually referred for psychiatric treatment.

The second group, Group C, consisted of individuals who were also self-referrals to the same university clinic. In the intake interview they also complained of intense and chronic anxiety, but the frequency of reports of depression was considerably less than with the group of psychoneurotics. The same type of clinical evaluation was accomplished with these subjects as with the psychoneurotics, and the final diagnosis was that of character disturbance, inadequate

personality. These subjects were also referred for psychiatric treatment.

The third group, Group X, consisted of volunteers for a study in physiological psychology. Before selecting each as a subject they were screened using an interview and the Minnesota Multiphasic Personality Inventory so as to avoid using any subject manifesting atypical amounts of anxiety or displaying symptoms of psychopathological conditions. Prior to the actual test situation the subjects were only told that they would be involved in an experiment involving physiology but were not aware of the actual procedure. The basic hypothesis of that particular experiment was that there would be a significant increase in blood sugar in a test situation where subjects received a randomly-administered electric shock at the volar forearm. The shocks were paired with buzzers according to a predetermined, random schedule. During the 37-minute test period four blood samples were taken using a lancet to puncture one of the finger tips of the left hand. A sample of .1 milliliter of blood was drawn into a graduated pipette and later analyzed for sugar content using a Lamotte Blood Sugar Kit. It seemed reasonable to postulate that if other studies had been correct, there would be a significant onset of anxiety during the test situation which would be verifiable through an increase in the blood sugar level. Each subject was asked to make a human figure drawing shortly after the second blood sample was taken, and it was assumed that if a significant degree of anxiety were present it would manifest itself somehow in this drawing. The fact that there was a significant increase, beyond the .01 level, in the blood sugar content for all subjects seems to be a reasonably sound indicator that the subjects were experiencing some anxiety or fear.

The fourth group, Group O, consisted of volunteers serving as controls in a Rorschach investigation. Each subject was asked to draw a human

figure before the administration of the Rorschach took place, and evaluation of the Rorschach data was made to insure that no distinct signs of psychopathology existed. These subjects were considered essentially normal.

A technique of quantifying line pressure, sketchiness, and shading was devised as a modification of the scoring procedure suggested by Caligor in his 8-Card Redrawing Test (1957). The scoring gradient for line pressure was based on the assignment of a numerical value of 1 to 5, with 1 being used for the light type of line and 5 being used for the heaviest type of line. A scoring gradient of 1 to 6 was constructed for degrees of sketchiness of lines, 1 being used for continuous solid lines, 2 being used with a continuous long but occasionally broken line, and so on, so that 6 was used for a highly sketchy line consisting of very short disconnected strokes. A scoring gradient of 0 through 3 was used for shading, 0 being given when no shading occurred and 3 when the shading was apparently to excess and quite heavy. In an effort to be reasonably detailed in the scoring, the drawings were subdivided into three sections so that one score was derived for the head, one score for the upper body (that being all above the waistline), and a third score for the lower body.

In addition to using scoring gradients for the line pressure, sketchiness and shading, scores of 0 or 1 were assigned for implied movement, profile drawings, presence of buttons, omission of feet, use of lower edge of the page as a baseline, and whenever the figure in the drawing was holding an object. The drawings were also measured for size. All of the drawings were coded so that when the scoring was accomplished it was not known which of the four groups the subject represented.

RESULTS

A statistical comparison of scorings for the four groups with regard to line pressure, sketchiness, and shading was accomplished by using the technique

of unmatched *t* test. Summaries of these data are found in Tables I, II, and III.

From examination of Table I it will be observed that Group C, character disturbances, differs significantly beyond the .01 level from all of the other groups in all three portions of the drawing. This indicates that these subjects consistently used lighter lines in their figure drawings than did any of the other subjects. There were no statistically significant differences between any of the other groups.

From examination of Table II it will be noted that Group N, psychoneurotics, shows a statistically significant difference from all other groups in the direction away from sketchiness.

In other words, neurotics were more prone to use continuous, unbroken lines. Data in Table II also indicate that Group C, character disturbances, used significantly more sketchiness than any of the other three groups.

Examination of Table III indicates that Group C, character disturbances, used significantly more shading than any of the other three groups and that Group X, experimental fear, used about as much shading as did the neurotics. The normals used the least shading of all, significantly less than either the character disturbances or experimental fear group but not significantly less than the psychoneurotics.

A comparison of the four groups on

TABLE I—A Comparison of the Groups for Line Pressure

Comparisons	Head			Upper Body			Lower Body		
	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>
N and C	3.75	2.55	3.86*	3.80	2.70	3.33*	3.65	2.45	4.46*
N and X	3.75	3.55	0.73	3.80	3.55	0.79	3.65	3.50	0.54
N and O	3.75	3.55	0.86	3.80	3.70	0.37	3.65	3.70	0.21
C and X	2.55	3.55	3.35*	2.70	3.55	2.85*	2.45	3.50	3.62*
C and O	2.55	3.55	3.83*	2.70	3.70	3.54*	2.45	3.70	5.06*
X and O	3.55	3.55	0.00	3.55	3.70	0.56	3.50	3.70	0.78

*significant beyond .01 level.

TABLE II—A Comparison of the Groups for Sketchiness

Comparisons	Head			Upper Body			Lower Body		
	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>
N and C	1.55	3.50	7.77*	1.40	4.40	12.59*	1.65	4.75	10.63*
N and X	1.55	3.00	5.90*	1.40	3.35	8.22*	1.65	3.40	6.78*
N and O	1.55	2.90	4.60*	1.40	3.05	6.60*	1.65	3.20	5.95*
C and X	3.50	3.00	2.13**	4.40	3.35	4.07*	4.75	3.40	4.93*
C and O	3.50	2.90	2.11**	4.40	3.05	5.00*	4.75	3.20	5.62*
X and O	3.00	2.90	0.36	3.35	3.05	1.11	3.40	3.20	0.83

*significant beyond .01 level.

**significant beyond .05 level.

TABLE III—A Comparison of the Groups for Shading

Comparisons	Head			Upper Body			Lower Body		
	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>	Respective Group Means		<i>t</i>
N and C	0.60	1.30	2.45**	0.85	1.80	3.65*	1.15	1.95	2.69**
N and X	0.60	0.90	1.36	0.85	1.40	2.22**	1.15	1.25	0.35
N and O	0.60	0.50	0.46	0.85	0.75	0.40	1.15	0.60	1.92
C and X	1.30	0.90	1.47	1.80	1.40	1.74	1.95	1.25	3.16*
C and O	1.30	0.50	2.99*	1.80	0.75	4.47*	1.95	0.60	5.92*
X and O	0.90	0.50	2.03**	1.40	0.75	2.94*	1.25	0.60	3.11*

*significant beyond .01 level.

**significant beyond .05 level.

all other variables was accomplished using the chi square technique. No significant differences were found with regard to the presence of buttons, omission of feet, or the use of the lower edge of the page as a baseline. Table IV lists three other variables in which statistical significance was found. These are inference of movement, the use of a profile, and the holding of an object in the hand of the figure.

Examination of Table IV indicates that the Group N, psychoneurotics, had significantly fewer drawings inferring movement. These data also indicate that Groups N and C, psychoneurotics and character disturbances, drew the figures in profile significantly more than either of the other two groups; and these same two groups, N and C, also drew their figures holding objects significantly more than did either of the other two groups.

A comparison of the four groups for size of drawings was made using an unmatched *t* technique. No significant differences were noted.

DISCUSSION

While this study began as a simple investigation into manifestations of anxiety, it would seem that the results go well beyond that particular consideration. It is difficult to state actually whether the first hypothesis formulated, that psychogenic anxiety would manifest itself in the line qualities of a figure drawing in a manner significantly different from essentially normal subjects, has been proven or disproven. Certainly there are distinct

differences in the figure drawings of the psychoneurotic and character disturbances when compared with both the experimentally induced fear group and the control group of normals. It might be postulated that some of these differences are the result of the pathological anxiety, yet it is just as easy to postulate that the differences in the drawings might result from the totalness of the pathology rather than a specific symptom in itself, especially since the drawings of these two groups are also significantly different from each other. The second hypothesis, that subjects experiencing experimentally-induced fear would manifest this fear in the line qualities of human figure drawings in a manner similar to psychoneurotic and character disturbance subjects, must be rejected. It is evident that the group manifesting experimentally-induced fear exhibited considerable difference from both Groups C and N and very little difference from the normal control group. In fact, they were essentially the same types of drawings as made by normals with regard to most of the variables measured. The only difference appears in the area of shading which might be considered as a key element in the interpretation of anxiety responses. Interestingly enough, the experimentally-induced fear group exhibited more shading than the neurotic group and almost as much as did the character disorder group, the drawings of which were by far the most heavily shaded. Thus it could be that shading in itself does indicate some focalization of anxiety.

TABLE IV--A Comparison of the Groups for Implied Movement, Use of Profile, and Holding Objects

Comparisons	Implied Movement			Profile			Holding Objects		
	Frequency		Chi ²	Frequency		Chi ²	Frequency		Chi ²
N and C	4	12	5.10**	9	10	0.00	11	12	0.02
N and X	4	10	2.89**	9	2	4.51**	11	2	7.29*
N and O	4	15	10.03*	9	2	4.51**	11	4	3.84**
C and X	12	10	0.10	10	2	5.93*	12	2	8.13*
C and O	12	15	0.46	10	2	5.93*	12	4	4.26**
X and O	10	15	1.71	2	2	0.00	2	4	0.08

*significant beyond .01 level.

**significant beyond .05 level.

Going beyond the considerations of anxiety, there are many distinct differences between the various groups which require consideration. The drawings of the neurotic might best be described as having a continuous line of reasonably heavy pressure, avoiding the use of sketchiness, and having only a modest degree of shading. The neurotic typically does not infer movement in his drawing, often uses the profile rather than full-face figure, and has a tendency to draw the figure holding some object. When comparing the drawing of the neurotic to that of the normal, we find that they are very similar in terms of line pressure and shading and are less sketchy than the normal.

The drawing of the character-disturbance subject seems distinctly different from that of the neurotic and distinctly different from that of the normal. It might best be described as a figure constructed with relatively light line pressure in which much sketchiness and shading are used. There is as much tendency to infer movement as is demonstrated in the drawings of normal subjects but also as much tendency to use a profile rather than a full-faced figure and provide the figure with an object on which to hold as does the neurotic.

Even though these are relatively small groups, the statistically-significant differences are somewhat impressive and do seem to lend support to the usefulness of the Draw-A-Person technique as a clinical instrument, not necessarily one from which absolute diagnostic information may be derived, but instead one which does yield some general diagnostic indications and thus becomes useful in screening situations. It should be emphasized that one of the groups used in this study was comprised of non-psychotic, non-neurotic, non-organic persons who still exhibited psychogenic anxiety and who were classified as character disturbances, inadequate personalities. In reviewing the literature pertaining to the Draw-A-Person Test, and for that

matter most all of the projective devices, it is difficult to find much reference to this relatively large group of people. Typically these are persons who are not incapable of functioning in their everyday environments; in fact, the majority of such persons probably lead relatively useful lives, carrying out occupational duties which require little decision making and which provide sources for dependency under stress conditions. It might be speculated that, because these types of individuals do not ordinarily manifest classic signs of psychoneuroticism or other major categories of psychopathology, they are often considered as "essentially normal" and could, accidentally, be included sometimes in control groups of studies investigating the diagnostic usefulness of test techniques. As the data presented here indicate, these persons do not perform as do "clinically appraised" normals, nor do they perform as do psychoneurotics; yet there is enough overlap between these three groups with regard to many of the variables tested to suggest that they could cause a confusion of the statistical significance of the data and thus a confusion in interpretation had they not been evaluated using a rather exhaustive clinical test battery.

It is appropriate to point out that while each group did manifest certain characteristics of the drawings which made that group significantly different from one or more of the other groups no single characteristic of any of the groups was so grossly different than all of the other groups that it could be specified as a highly valid diagnostic indicator. On the contrary, certain subjects of each group exhibited characteristics in their drawings which were more like those in another group than in their own group, and all of the groups appear to have certain common characteristics to all of their drawings. This factor seems to give strong support to the idea that human figure drawings, or for that matter most any projective technique, should

be interpreted from a global point of view wherein consideration is given to the entire data collected and with particular reference to the inter-relation of the parts of the data rather than attempting to isolate specific variables such as line pressure or shading and attach special meaning to these variables.

SUMMARY

Human figure drawings were collected from four groups of 20 subjects each. The groups consisted of psychoneurotics, character disturbances, normals, and subjects experiencing experimentally-induced fear, all of whom were evaluated using typical diagnostic clinical procedures. The drawings were scored for ten variables and significant differences were discovered in six categories which differentiate the psychopathological groups from each

other and from the other two groups. The drawings of subjects experiencing experimentally-induced fear showed little difference from the drawings of normals except with regard to an emphasis on shading. It was concluded that the Machover Draw-A-Person Technique does have clinical usefulness as a screening device but that no single element of the drawing can be regarded as a conclusive diagnostic indicator.

REFERENCES

- Machover, Karen. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Charles C. Thomas, 1949.
- Sundberg, N. D. The practice of psychological testing in clinical services of the United States. *Amer. Psychologist*, 1961, 16, 79-83.
- Swensen, C. H. Empirical evaluations of human figure drawings. *Psychol. Bull.*, 1957 54, 431-466.

Received April 26, 1962

Some Personality Correlates of Styles of Interpersonal Thought

ROBERT H. KNAPP, HERBERT GEWIRTZ, AND JULES D. HOLZBERG
Wesleyan University and Connecticut Valley Hospital

In the latter part of the 19th Century, Brentano (Boring, 1950) undertook to formulate a so-called "act" psychology, building upon the fundamental Aristotelian tenets concerning the nature of psychic processes. In effect, Brentano argued that the thought process consists of a person engaged in a psychic act, i.e., knowing, loving, hating, etc., with respect to some palpable object. In the present paper we are presenting some results obtained by a technique which derives from Brentano's basic conceptions of the thought process. Our concern is with such patterns as they apply to human diadic relations. We have sought, therefore, to explore responses to a range of human thoughts circumscribed by the fact that the subject and object must be personal pronouns, and that the acts associating them be among the most prototypical.

In short, we have evolved an inventory of thoughts whose construction will presently be described in greater detail. This inventory of thoughts concerns the relations of two persons according to the general schema set forth by Brentano in his "act" psychology.

It would appear that Freud fairly early in his career recognized the validity of the proposition that many thoughts, especially those relating humans, follow the pattern of subject, act, and object. This awareness is most notably illustrated in his study of homosexuality in the Schreiber case (Freud, 1953). Here Freud demonstrates that the thought "I love him" underwent a series of transmutations calculated to disguise the homosexual motive. Thus, we may note in passing that Freud recognized the importance of the basic diadic proposition, at the

same time observing that defense mechanisms are intimately involved in the manipulation and distortion of such propositions. In passing, we may say that it remains our hope to employ the Inventory we have devised for that purpose of identifying styles of character defense. This object, however, lies beyond the scope of the present paper.

No introduction to the general subject of diadic thought propositions as we have described them would be complete without mention of the great importance attached to interpersonal relations as the shaper of thought processes by Sullivan (1947). Nor should we fail to note the now classical writings of George Mead (Strauss, 1956) identifying diadic perception as an absolute requisite to the development of representational speech.

Utilizing this conception of thought processes with regard to diadic relationships, it is the purpose of this study to explore the ways in which an individual orders differing diadic propositions and in turn relate these orderings to certain personality characteristics.

METHOD AND PROCEDURE¹

The subjects for this investigation were 60 male sophomores attending Wesleyan University, who volunteered and received a small stipend for participation in the study.

The personality correlates used as independent variables in this study were derived from six psychological instruments that had been administered to the subjects approximately one year earlier at the beginning of their freshman year. From these six

¹The collection of the data reported in this article was accomplished by David Parker, graduate student at Wesleyan University.

instruments a total of 33 personality variables were selected, as follows:

1-10. Ten variables from the Minnesota Multiphasic Personality Inventory (MMPI) consisting of the nine psychopathological scales and the scale of social introversion.

11-16. The six primary variables of the Allport Vernon Scale of Personal Values (A-V), namely, measures of theoretical, economic, aesthetic, social, political, and religious interests.

17. The score on the Terman Concept Mastery Test, measuring superior intellectual talent.

18-19. Two variables drawn from the Strong Vocational Interest Inventory, the M/F Scale and the Scale of Scientific Interest.

20-27. Eight variables from the Myers-Briggs Personality Test (M-B), based on the Jungian typological scheme and measuring extroversion-introversion, sensation-intuition, thinking-feeling and judging-perceiving.

28-33. Six measures derived from the Knapp Metaphor Scales (KMS), described in an earlier publication (1960), and measuring subjective interpretation of time, conscience, death, success, self-image, and love.

Thus, there were a total of 33 personality measures which for the purpose of this investigation served as our independent, or external, variables.

The dependent variables in the present study have been derived from an instrument developed by the authors. This instrument we have called the Interpersonal Transactional Inventory (ITI). In our subsequent discussion we shall review the manner in which this Inventory was evolved. Here, we shall merely describe its structure and its mode of administration.

The purpose of the Inventory is to present to the subject a catalog of propositions which involve all possible combinations of first person, second person, third person masculine,

and third person feminine pronouns in both subject and object, linked by one of four, as we conceive them, primary acts or verbs. These four verbs are "like," "dislike," "know," and "don't know." The first two are affective in character, the second two cognitive. The first and third are affirmative or positive, and the second and fourth denying or negative. We shall term our first and second person pronouns hereafter as "intimate," the third person pronoun as "distantiate." It should be said in passing that in preliminary experiments not reported here we sought to include the acts of love and hate, and the acts of trust and doubt. We finally reverted to the simple schema of the four acts described above in the interests of economy and simplicity. It will be seen that the possible combinations of the four persons in the subject and the four persons in the object with our four acts will lead to the emergence of 64 "interpersonal" propositions, each of which was typed on a 3 x 5 card.

The administration of the Inventory was accomplished as follows: Four sets of 28 cards each were assembled. The first of these contained all propositions in which the first person appeared in either subject or object; the second set, all propositions involving the second person; the third set, all propositions involving the third person masculine; and the fourth set, all propositions involving the third person feminine. Thus, all statements excepting the reflexives (same person, subject and object) appeared twice among the cards comprising the four decks.

Each subject was required to sort each set into seven stacks of four propositions each, according to "the frequency with which these thoughts are harbored by your fellow students at Wesleyan." These instructions were designed to serve as a projective basis for getting at personal preferences for these thoughts. After the four stacks had been sorted, it proved a simple

matter to score each proposition for its position on the seven-point scale defining the frequency with which the proposition was judged to be harbored.

In the present instance we have confined ourselves to a fairly simple scoring method in which we have obviously neglected some of the more subtle interactions potentially available in this test. The sixteen dependent variables deriving from this Inventory comprise the following:

1. Preference for the act "Like"
2. Preference for the act "Dislike"
3. Preference for the act "Know"
4. Preference for the act "Don't Know"

5. Preference for affective acts (Like and Dislike)

6. Preference for cognitive acts (Know and Don't Know)

7-10. Preferences for 1st, 2nd, 3rd masculine, 3rd feminine persons as subjects, respectively

11. Preference for (intimate) subjects (1st and 2nd persons)

12-15. Preference for 1st, 2nd, 3rd masculine, 3rd feminine persons as objects, respectively

16. Preference for intimate objects (1st and 2nd persons)

These sixteen measures were obtained for each of our subjects and thereafter a correlation matrix was computed relating our 16 dependent variables and our 33 independent variables, earlier described. This 49 x 49 matrix was next factor analyzed. Eight factors were extracted and rotated according to their Varimax procedure.

RESULTS

Of the eight factors, four clearly involved the joint participation of both the dependent and independent variables with a high order of confidence. Two more factors involved our dependent variables in a strongly suggestive degree, and the remaining two factors involved only independent variables.

It will be seen that the most important dimension of Factor I is de-

rived from the Myers-Briggs sensation-intuition scale. With intuition are associated aesthetic interests as measured by the Allport-Vernon, femininity as

FACTOR I

Variable	Loadings	
	Positive	Negative
You (obj.) (ITI)	+.34	
Conscience (KMS)		-.31
Success (KMS)		-.38
Self Image (KMS)		-.61
Love (KMS)		-.43
M-F (MMPI)	+.35	
Economic (A-V)		-.45
Aesthetic (A-V)	+.49	
Science (Strong)	+.36	
Sensation (M-B)		-.75
Intuition (M-B)	+.77	

measured by the Minnesota Multiphasic Personality Inventory, scientific interests as measured by the Strong Vocational Interest Blank, and the choice of propositions involving "you" as object. At the opposite polarity of this factor, and associated with sensation on the Myers-Briggs, we find high economic interest on the Allport-Vernon, buoyant images of Self, Love, Success, and Conscience on the Knapp Metaphor Scales. The participation of our Inventory in this factor cannot be confidently interpreted.

Factor II is of considerable interest insofar as it relates scales of the MMPI to our Inventory. It will be observed that substantial positive loadings are found on the schizophrenic and paranoid scales as well as on the choice of intimate persons as objects. In contrast, the choice of

FACTOR II

Variable	Loadings	
	Positive	Negative
Me (obj.) (ITI)	+.70	
You (obj.) (ITI)	+.46	
Him (obj.) (ITI)		-.41
Her (obj.) (ITI)		-.64
Me and You (ITI)	+.80	
Paranoia (MMPI)	+.48	
Schizophrenia (MMPI)	+.41	

distantiate persons as objects carries very high negative loadings.

Factor III consists entirely of inde-

pendent variables, and is dominated

FACTOR III

Variable	Loadings	
	Positive	Negative
Depression (MMPI)		-.67
M-F (MMPI)		-.39
Psychasthenia (MMPI)		-.41
Social I. E. (MMPI)		-.81
Religion (A-V)	+.35	
Terman		-.45
Extrovert (M-B)	+.85	
Introvert (M-B)		-.87
Feeling (M-B)		-.34
Judging (M-B)	+.47	
Perceiving (M-B)		-.48

by the dimension of introversion-extraversion of the Myers-Briggs. With extraversion, we find associated the judging scale of the Myers-Briggs and the religious scale from the Allport-Vernon. Associated with introversion are four of the MMPI scales, notably the scale of social introversion and the qualities of feeling and perceiving. Also included here is the Terman Concept Mastery Test with its high premium on verbal skills. This factor, which holds less interest for us than the others, can probably be described as an introversion-extraversion factor, with components of neuroticism positively associated with introversion.

Factor IV is of major importance with respect to our Interpersonal Transactional Inventory, with particular focus on choice of different "acts." Here we find a major dichotomy established between affirmative cognition (know) and negative affection (dislike). This dichotomy is factorially related with the Myers-Briggs scale of judging vs. perceiving. Associated with positive cognition to a suggestive degree is hypomania from the MMPI.

FACTOR IV

Variable	Loadings	
	Positive	Negative
Dislike (ITI)		-.78
Know (ITI)	+.87	
Like and Dislike (ITI)		-.69
Like and Know (ITI)	+.71	
Hypomania (MMPI)	+.34	
Judging (B-M)		-.56
Perceiving (B-M)	+.58	

Factor V is dominated by the opposition of intimate vs. distantiate subjects. Thus, choice of "I" and "you" as the subject is opposed to preference for "he" and "she." The former subjects are positively associated with the political scale of the Allport-Vernon and the latter with the Science scale of the Strong Vocational Interest Inventory.

FACTOR V

Variable	Loadings	
	Positive	Negative
I (subj.) (ITI)		-.67
He (ITI)	+.31	
She (ITI)	+.61	
I and You (ITI)		-.87
Political (A-V)		-.39
Science (Strong)	+.48	

Factor VI involves only independent variables. It is dominated by the thinking-feeling dimension of the Myers-Briggs. With feeling are associated religious values from the Allport-Vernon, effeminacy from the MMPI and a buoyant image of conscience from the Knapp Metaphor Scales. Associated at the opposite pole

FACTOR VI

Variable	Loadings	
	Positive	Negative
Conscience (KMS)	+.33	
M-F (MMPI)	+.32	
Theoretical (A-V)		-.37
Economic (A-V)		-.52
Political (A-V)		-.31
Religious (A-V)	+.58	
M-F (Strong)		-.58
Thinking (M-B)		-.66
Feeling (M-B)	+.71	

with thinking and masculinity from the Strong are the theoretical, economic, and political scales of the Allport-Vernon. This factor may be described as a sort of tender-mindedness vs. tough-mindedness, but it is of less importance to us since our dependent variables are not involved..

Factor VII is of interest since it involves the problem of masculinity-femininity and neuroticism in so striking a degree. Here we find negatively

associated most of the clinical MMPI

FACTOR VII

Variable	Loadings	
	Positive	Negative
She (subj.) (ITI)		-.34
Hypochondria (MMPI)	+.66	
Depression (MMPI)	+.41	
Hysteria (MMPI)	+.68	
Psychopathic		
Deviate (MMPI)	+.61	
M-F (MMPI)	+.52	
Paranoia (MMPI)	+.44	
Psychasthenia (MMPI)	+.64	
Schizophrenia (MMPI)	+.66	
Hypomania (MMPI)	+.46	
M-F (Strong)		-.44

scales and the tendency to choose on our Inventory the third person feminine as subject. Opposing this is the masculinity scale of the Strong Vocational Inventory as the other variable with a significant negative loading. It is strongly suggested that preference for the feminine subject is negatively associated with general neuroticism and, in particular, effeminacy.

Factor VIII is related in an interesting manner to Factor IV. In Factor IV we saw the opposition of positive cognition to negative affection. Here we find the opposition between positive affection and negative cognition. The suggestive association of affirmative images of success and the social value scale from the Allport-Vernon with negative cognition does not admit of immediate interpretation and may, indeed, be a statistical artifact.

FACTOR VIII

Variable	Loadings	
	Positive	Negative
Like (ITI)	+.65	
Don't Know (ITI)		-.75
Like and Dislike (ITI)	+.52	
Like and Know (ITI)	+.51	
Success (KMS)		-.31
Social (A-V)		-.34

DISCUSSION

The results of the factor analysis indicate that the three elements of our interpersonal transactions (subject, verb, object) are meaningfully

related to a number of our external variables. All of the subjects participate in Factor V. All of the objects participate in Factor II. The interactional verbs are found in Factors IV and VIII. These findings would support the factorial validity of our Inventory.

No definitive theoretical interpretation of the factors found is possible at this time, but it should be noted that a number of them clearly suggest some interesting possibilities. Factor II reveals that intimate objects, particularly "me," are positively related to paranoia and schizophrenia, lending confirmation to the clinical observation of the paranoid individual being preoccupied with acts directed toward the self. His is the feeling of being the object of external forces and many of his ideas of reference center on this conviction.

Factor IV relates two cognitive processes—"know" on the ITI and a preference for perception on the Myers-Briggs. The inverse relationship between negative affection (dislike) and hypomania seems to mirror the classical dynamics of the manic who is denying his hostile impulses. The interesting relationship between negative affection (dislike) and a preference for judging (moral fervor) on the Myers-Briggs should also be noted.

Factor V relates political interest to a preference for intimate subjects, particularly "I," as the initiators of action while scientific interest is related to more distant subjects (he and she). The difference in the temperamental qualities of the scientist and politician is thus strikingly supported by the ITI.

Factor VII describes an inverse relationship between "she" as the initiator of action and many of the MMPI scales. In this male college student body, it may be assumed that thinking of the feminine sex would be negatively related to pathology. It should also be noted that thinking of "she"

is positively related to masculinity on the Strong

A number of directions for continued research in this area have been considered by the authors. One lies in the study of sex differences that would attempt to detect differences in thinking styles between men and women. A number of the factors clearly suggest a relationship between performance on the Interpersonal Transactional Inventory and a number of pathological indicators. Research might be directed not merely toward differentiating the normal from the pathological, in general, but also toward differentiating between various pathological groups. Still another area of investigation might be the study of the relationship of styles of thought to other modes of behavior, e.g., perceptual processes, preferences for sensory modalities, etc. Perhaps most interesting of all is the possibility that the Interpersonal Transactional Inventory may be adapted for the study of styles of personality defense. As we have noted earlier, it was in this context that Freud first discussed kindred interpersonal propositions in connection with the Schreiber case. The technique of such an inquiry cannot be explored here, except that, in principle, we would seek to determine the degree to which certain propositions are the object of sensitization or defense.

It should be stressed in conclusion that the scoring which has been devised for the current investigation is rather crude. For example, we have not considered interaction scoring categories involving subject and act, subject and object, or act and object. However, we have used the Inventory as a clinical instrument by examining individual protocols in enough in-

stances to know that analysis in terms of such interactions can be most illuminating of the total processes of the individual. Probably future developments of the Inventory should admit more subtle and exhaustive methods of scoring. We have sought here to set forth only the theory of the Inventory and a demonstration that reactions to it are meaningfully related to a number of established psychological measures.

SUMMARY

This paper describes the rationale and development of an Interpersonal Transactional Inventory consisting of diadic propositions of four pronouns as subject and object (first, second, third person masculine, third person feminine) and four acts or verbs linking subject and object in an interpersonal transaction. Measures derived from the Inventory were factorially related to a number of independent variables derived from standard and previously published personality scales. The compositions of the factors derived from factor analysis support the factorial validity of the Interpersonal Transactional Inventory.

REFERENCES

- Boring, E. G. *A history of experimental psychology*. New York: Appleton-Century-Crofts, 1950.
- Freud, S. Psycho-analytic notes upon an autobiographical account of a case of paranoia. In S. Freud, *Collected papers* (Vol. III). London: Hogarth, 1953.
- Knapp, R. H. A study of the metaphor. *J. proj. Tech.*, 1960, 24, 389-395.
- Strauss, Anselm (Ed.). *Social psychology of George H. Mead*. Chicago: University of Chicago Press, 1956.
- Sullivan, H. S. *Conceptions of modern psychiatry*. Washington, D.C.: William Alanson White Foundation, 1947.

Received April 16, 1962.

Certain Theoretical Considerations In Relation To Borderline Schizophrenia and the Rorschach

ROBERT S. McCULLY¹

Cornell University Medical College, and The New York Hospital
(Payne Whitney Psychiatric Clinic)

Rorschach records which reflect a so-called borderline schizophrenic structure sometimes constitute a very practical problem in communication for the Rorschach worker. Much has been written describing borderline schizophrenic records, but there has been little focus on theoretical considerations about this kind of picture. These records may add to our knowledge of the nature of the schizophrenic condition. Considering the nature of the process of what may be taking place in the Rorschach examination of these patients may be fruitful. The purpose of this paper is to illustrate what might be called "process analysis" in a particular borderline Rorschach record and to relate this process to a theoretical framework applicable to this kind of condition.

In borderline cases, it is not infrequent to find some disparity between behavioral symptoms and the extent of apparent psychopathology in the Rorschach. A Rorschach worker may know exactly what another worker is talking about when he describes the Rorschach qualities of a borderline case, but individuals from other disciplines may not. Not infrequently, the Rorschach worker hears, "but there are no clinical psychotic symptoms, what does it mean when you say the Rorschach picture has schizophrenic features?" Too often, in such a seemingly contradictory context, Rorschach data may be dismissed for one reason or another.

David Shapiro (1954), in his cogent treatment of the problems surrounding borderline psychotics, has pointed

out how rigid diagnostic demands can cloud the issue. However, a flexible set of diagnostic criteria does not necessarily imply a broader range of understanding about the condition being described. We need to look squarely at just what we may be describing and at the conditions that surround it. The value ascribed to data should not be dependent upon what can be demonstrated "clinically," though they may have different import. Valuable data about the psyche of the individual may be discarded in paying homage to the external situation. Something of the paradox of the scientific bias may obtain here. Peter may be being robbed to pay Paul.

This whole matter requires the consideration of what is meant by "latency." In Rorschach, conflict, symbol, dynamic facets, and other constellations in the content material may reflect latent qualities, but are peculiar logic, bizarre expression, and other anomalies (such as may be associated with the *process* of what is happening while taking the Rorschach) "latent" anymore, since they have occurred? When one considers what has been produced, for the moment laying aside the nature of the stimulus, there is nothing inherently more mysterious or deep about peculiar associations given to ambiguous stimuli than to anything else. Peculiar associations are not in themselves material from a deep layer, they have *happened*. In the best scientific tradition, *something* has been demonstrated; data have formed a tangible pattern. Their relation to "deep layers" is another issue altogether. When considering latency in relation to Rorschach material and the behavior of the individual, perhaps one should not ask,

¹ The author wishes to express his appreciation to Dr. Florence Miale for suggestions and comments.

is this bizarre material "latent," but what was the nature of the stimulus when this kind of data came into being or no longer remained "latent"? The nature of the stimulus is, of course, the crucial difference between what may be elicited clinically and what may be elicited in Rorschach. That findings from both often coincide is more a testament to the communality of human experience than to similarity in method. This distinction has been overly simplified for purposes of clarity. Even though Rorschach blots are relatively unstructured, this does not mean they are therefore removed from life. The subject brings the experiences of life to the blots and makes life of them; we get to see something of what life means to him quite over and beyond the stimulus. Miale (1959) has pointed out that the extent of structure in the stimuli may not be so crucial as one may suppose. She remarked that the life situation is really quite as ambiguous as the Rorschach. The point we wish to make is that the process one finds going on in Rorschach represents in microcosm essential features going on in the macrocosm of living — even though specific behavioral patterns may seem contradictory.

In considering borderline cases and the Rorschach, Forer (1950) has pointed out how, in this method, the individual cannot relate new experiences to old ones, because of the highly ambiguous qualities of the situation. This, of course, is one of the reasons the Rorschach method offers us such rich and pertinent data about the individual. The soundness of the Rorschach subject's judgment hinges on his capacity to show appropriate behavior despite minimal cues. Forer indicated that thus, the individual reveals more of himself by having been thrown more on his internal cue system. Forer concluded that the amount of the stimulus structure which the individual requires to behave appro-

priately determines whether a disorder is overt or latent.

It should be understood then, that when one speaks of a thinking disorder on psychiatric examination, and when one speaks of a thinking disorder on Rorschach examination, one may be speaking of what has happened under quite different kinds of stimuli. As forms of stimuli, one focuses more on the degree of formal adaptation, and the other on what happens when the range of stimulation is extended and the frame of reference is shifted from the familiar to the unknown. One might expect that the "frankly psychotic" individual would show clear-cut evidence of the disturbance whether the range of stimulation be narrow or wide. Nevertheless, sometimes the frankly schizophrenic patient shows less evidence of disturbance within the Rorschach than does the borderline case. However, it is primarily in the "neither-nor" cases that the data from the two methods of examination may diverge.

In considering borderline cases, one can say that if the range of stimulation is extended, some people produce psychotic features, while others do not. This does not necessarily mean the one is psychotic and the other isn't. It means that there are people who are vulnerable to psychotic responses under certain conditions, and the difference between the presence or absence of such material lies in the nature of the psyche of the individual under consideration. Rorschach behavior is just as clinical as any other behavior; what then, of the matter of "correct" diagnosis? One must be practical, and lines must be drawn. Does one make a "correct" diagnosis when it is based exclusively on relatively well structured sets of stimuli? At least in some cases, "correct" diagnosis becomes rather like an academic exercise based on practical needs. Our need to classify what may be by its very nature ambiguous into something without ambiguity may de-

feat its own ends. Let us not stumble over our own egos.

That there are people who, under certain conditions only, tend to produce psychotic material, brings us back to the borderline schizophrenic. Perhaps Baynes (1940, p. 7) has given us the most graphic description of the borderline case:

"Outwardly, the borderline schizophrenic may have a trim, polished, impeccable mask, which attracts without promise of personal warmth. The smile is inclined to be 'grooved', lacking the spontaneity of feeling. At every point one may discover a contradiction between the outer aspect and the inner life. The expression shown to the world is non-committal and reticent, while the archaic emotional welter within is often volcanic and terrifying. The concealed affects are liable to be distorted and perversely exaggerated, because the subject's mind is flooded with archaic content which he cannot express . . . but, as a rule, he may have little knowledge of his own motives . . . At all costs he must appear like those around him: there must be no chink in his protective armor of simulation."

In such a patient one would rather *expect* to find differences between an externally oriented examination and an internally oriented one. When the rug is pulled out from under an individual with rigidly structured, but tenuous outer balance, one may expect a tumble. A conclusion based exclusively on what the patient does in either sphere (relation to the outer world versus relation to the inner world) may not be "correct" in either case. They both not only overlap but are interdependent. We must bridge the gap that the different diagnostic approaches to the matter may leave, and we must do it in such a way as to do justice to the state of the individual's global prospects.

Two concepts which support points raised in this paper need to be described before we take a look at the case materials. The first of these is a proposed theoretical frame against which we may judge spheres of movement in what may be referred to as the disturbance process (and corre-

sponding adaptive process) in the borderline schizophrenic. For the present purposes, this schema is meant to be kept very simple. One may conceive two planes, one horizontal and the other vertical (see Fig. 1). On the horizontal plane, withdrawal from (the other) designates the left end

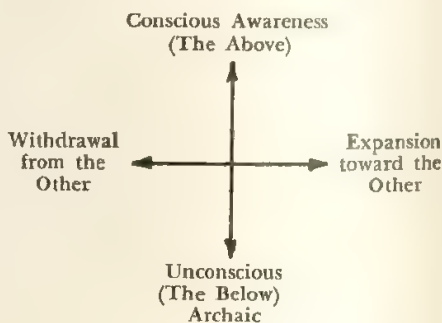


FIGURE 1

of the continuum, and expansion toward (the other) designates the right end of the horizontal plane. The other axis is the vertical one, with conscious awareness at the upper end and unconscious at the lower end. The movement in the disturbance (or adaptive) process may occur anywhere along these axes, and we will be interested in where the emphasis or focus may appear. The focus of interpretation in psychiatric examination tends to be along the horizontal plane, while in the Rorschach examination the focus is most often along the vertical. This may be more a matter of the emphasis taken by the interpreter than extent of focus in each method per se. That both techniques overlap and that both planes may be involved is not the point here. The point is that these two directions offer us a frame against which we may describe what is going on in a borderline case. The axes will be familiar to those who know handwriting theory². Analogy will be made to Rorschach cards and these planes, both in terms

² As a matter of fact, the patient to be described showed breaks in logic only on the vertical axis in his handwriting sample.

of card position and within the cards themselves.

The second concept may be called "process analysis" and it is related to what has been described above. The distinguishing feature of this kind of Rorschach analysis is its focus. It should not be viewed as a new way of analyzing Rorschach material. Rather, it is a particular kind of focus in the analysis of sequence through the Rorschach. It is a result of asking, what process is going on in the patient at this point? To be sure, it overlaps with content analysis, but more especially, it is an attempt to isolate and interpret the process the patient uses in structuring and integrating "whatever comes to consciousness." To accomplish this, one must make defenses, dynamics, and other ordinary foci in Rorschach interpretation peripheral for the moment. The focus here is on what the individual is doing at any one time and its relation to the whole, a subtle use of symbol in the realm of what the patient is doing. It is not independent of any other Rorschach feature (dynamics, defenses, symbol, intellection), but it is an effort to draw attention to the qualities of what the individual experiences as he goes through Rorschach material, and to interpret these experiences as symbolic material. This process is best understood without using defense terminology, since these terms may obscure just what we want to know into stereotypes.

An illustration of looking at the process in an actual Rorschach image may help clarify what is meant by this term. The material is taken from Schafer's (1954, p. 365) examples of cases showing obsessive-compulsive defense operations. The responses were given by a male patient, 43 years old, to Rorschach Card V:

1. "I started off to have an idea but it's hard to carry out. I see the back half of animals like deer or antelope (side wing and projections), but it's hard to tell what happened to their front ends. They get lost in all this stuff (middle D). The

back legs, the tail and the back part of the back are pretty clear."

2. "These are large, swept-back wings here—looking at the black. Clumsy wings; they wouldn't work well, I suspect. Tying in with the wings perhaps are a head like that of a vampire bat with very large ears extending away from me and that would tie in with the legs toward me. It would be a pretty punk sort of a bat, I expect."

Schafer's interpretation of these responses may be summarized by emphasis on the anal-oral contrasts, minimization, and the rigidly defensive handling of the aggressive potential in the subject's images. In looking at this material on a process level, we simply change the focus, and this in no way challenges the validity of Schafer's interpretation.

The subject is having an unjustified struggle with the card that offers an easy, conventional form-percept. The least complicated way to make a connection with the horizontal plane (the other) is a direct one. The subject imposes barriers between himself and the conventional, suggesting no easy access to the horizontal. When he attempts the unique, which is to say, when he follows a prompting to be different or unconventional, it fails to come off. There is such a gap between the lofty expectation of the self and the "disorder" in the reality of the blot with which he has to cope. One suspects a quality of disconnection between his own ideas; the front gets lost in the middle and in the problem with the center. His choice of vehicle in his effort is from the instinctive world, the animal, that which is symbolic of "the below." He verbalizes the clarity with which he experiences the outer parts of his image (in the horizontal sphere of the blot) and the lack of clarity in the central portion (vertical). His comments may refer to objective aspects of the blot qualities, but we must remember that this concern developed out of his own choice. It was the subject who selected these blot qualities on which to focus.

The need for *order* appears both in regard to the heads of his animals (instinctive order), and in regard to the quality of the center (hazy) of the blot itself. He shifts responsibility for the dilemma to the blot and away from his own first urge to be "different." The subject has a problem about his instinctive world in relation to both axes in our schema. One may expect that the subject experiences vague intrusions from "the below" just at the time he is the most concerned with his relation to the outer-other. Yet, it is not so much *relation* that he is concerned with as it is how he appears and his showmanship. This is where he is vulnerable. Generally, his judgment is good; when it diminishes, he knows it, but he tends to project the responsibility.

In his second response, he gets around to coping with the ordinary or conventional. How does he do this? He at once becomes super-critical and condescending. We see that his ambition has such large wings — the inflating and the inevitable failure because of the inflated expectation. He virtually pulls the conventional apart and puts it back together again, as if to see what makes it tick. Such unnecessary gyrations with a simple percept. The fact is, he can't be content with the simple and he has to make it complex; in part, his self esteem depends on this process. There is a quality of glorifying and denigrating the vampire bat. He finds an excuse for being nasty through a self-depreciatory attitude. He makes a literal separation between the vertical and horizontal aspects of the blot, and then he ties them together again. He operates, then he sutures. He can do this much more successfully (perhaps in spite of himself) when he is dealing with the ordinary than when he is "being unusual or different." He simply can't bear being ordinary, and he segmentalized and analyzed the popular bat into a poor specimen indeed. This is a subject who is uneasy with himself and uneasy with the other.

His need to be different lures him into the deep waters of the below and the unknown. However, he admits it when he gets "out of his depth," though he rejects the responsibility for it. He always concerns himself with a reference to the horizontal, though it is more a concern with the impression he is making on the other than a mature concern with relation to the other.

Table I contains the Rorschach materials of a twenty-three year old, single, male subject (WAIS verbal I. Q. 113). We shall now look at his responses and comments as they appeared in sequence, but with a focus on the process going on in the subject and its relation to the axes described. This should not be considered as an exhaustive Rorschach analysis, but an interpretation which stresses certain points. A word of caution obtains with regard to this interpretation. It should be remembered that this is a subject who is relatively open to the symbolic and awareness of his own processes. His concern with the position and plane of the Rorschach cards was by no means accidental or commonplace. The position of the cards had a special meaning for him. This is an example of what Bleuler (1926, p. 31) meant when he described affect as a director or selector of attention. We do not know what parallel obtains between this subject's turning the cards and what may be going on when other kinds of Rorschach subjects turn the cards.

INTERPRETATION

We will follow the process of Rorschach events from the subject's first remark to his last. In some instances, smoothness may be sacrificed to capture the essence of the process. It is essential to keep the features of Fig. 1 in mind to follow the interpretation.

How does the subject begin a new experience when the cues are up to himself? He begins by whistling in the dark and repeating, "they are ink-blots" (he had been informed how

TABLE I—Rorschach

I.

Do you want the first impression?
It looks like an ink blot (he had
never seen the cards before).
Dare I pick out of it a piece
(to himself)? This one isn't
particularly disturbing. First I
see the whole thing, then two
parts.

- | | | |
|-------------|--|---|
| 1. \wedge | 1'39". Sea horses come to mind,
but they don't have tails.
W F A | 1. The shape (includes the two side Ds). |
| 2. \wedge | Really it looks like a crest of
some sort.
W F Shield | 2. A shield, It's shaped like that. |
| | \vee Can I turn it upside down, it is
much better, ah! | |
| 3. \vee | It looks like a mask.
WS F Mask | 3. The shape. |
| 4. \vee | A cow with a crown, these are
the horns.
WS F (FC') Ad O | 4. This is the forehead (lower space area),
these are the horns jutting out, and all
this above is the crown. |
| 5. \wedge | A cat with his tongue out.
WS F Ad | 5. The shape, eyes, mouth and the tongue
out (lower central d). |

II.

\wedge You don't care really which
way? $\wedge \vee \wedge$ It looks messy,
therefore it may be something
physiological. All of them are
symmetrical.

- | | | |
|----|--|---|
| 1. | \wedge 78". Something biological, a dia-
gram.
W CF AT | 1. Something of that shape. (?) Yes, like a
diagram, the colors. |
| 2. | $>$ Ah! Sputnik breaking through
outer barriers into space. <i>I do
much better with it on its side.</i>
The red could signify the flash
of rocket fire.
SD Fm (CF) Rocket | 2. This white part, rapid speed and the
blast. |

III.

- | | | |
|-------------|--|---|
| 1. \wedge | 3". Very good! Two cannibals
over a pot, that looks much
better, and they are cannibals
because the red is a bloody ob-
ject being cooked $\vee <$. Two
very strong (he means the per-
cept is strong) human figures, I
suppose even women at that
$< \vee$
W M (CF, FC') HP | 1. Shape and movement. They are two
women. (?) The simulation of breasts. (?)
Some sort of red object, bloody meal. (?)
This is a black kettle here. |
|-------------|--|---|

IV.

1. \wedge 3". A lion's rug on the floor stretched out and comfortable. The head is up at the top, the feet are stretched out. Some sort of skin, animal skin on a wall or something, a lion or something, not a small animal.
W F→Fc(?) Aobj.
1. A tanned hide, the shape. (?) Well, just the shape. (?) I didn't notice the quality of any of these. (?) inside or outside?) We are looking at the inside. That's the way you hang them, the hair is the other side turned to the wall to protect it. This is the tanned leather side.

V.

1. $\wedge \vee \wedge$ 7". An insect of some sort flying. \vee I may think of biology cards for all of them.
W F→FM A P
1. Shape. (?) It looks insectual. (?) I don't see it in any motion.

VI.

1. \wedge 7". A hide tanned, that's all I can think of. Am probably afraid of anything realistic on these things, just descriptive.
W Fc Aobj. P
1. The shape. It looks less like a hide than the other one. It's a feline specimen with whiskers and a long, long nose. Because of the quality of this, I would turn it the other way! Uck! It's pretty awful, hairy!
2. $>$ An open mouth, on the far left, maybe of a whale, the teeth there.
ds F Ad
2. Big and prehistoric looking, those look like teeth.

VII.

1. $\wedge \vee > <$ 39". Aha! Two women facing in opposite directions with black fur bear hats on, in a chorus line, dancing.
W M (FC', Fc?) H
1. The shape looks like just that, dancing too. (?) fur hat?) Well, the shape, hair couldn't extend so far and it's dark.
2. \wedge These balancing rocks, they would have to be pretty well balanced to be symmetrical.
W mF N
2. The shape. (?) Just the shape. The balance, even though it stays there without really being in balance.

VIII.

1. \wedge Colors no less. $\wedge \vee < \wedge$ 63". Two dogs, don't know what they are doing, pink at that. $>$ Dogs looking at a lake with the image reflected back. Never saw pink dogs before.
D F/C A P
1. Shape, and reflection. The color has nothing to do with it. I couldn't explain the color, maybe it's sunrise.

IX.

- $\wedge \vee > <$ This one has me . . . $> \wedge$ It doesn't inspire me . . . the colors don't mean anything. $> \vee <$ You know, I prefer to look at all of them in a longitudinal way $>$. I'm disappointed in my imagination.
1. \wedge 3'50". Some sort of sea life maybe, the early phyla, it turns up its mouth to get food.
W CF N
1. Early plant life, plant-animal, what's the name? Way down in phylum of things. They turn down from above like this, and they have this green on them and they take in food.

X.

1. Δ 74". A real fantasia, something in a children's book, my first preoccupation is with color.
W CF Art
2. Δ Two men from Mars fighting it out, each has a knife in his hand.
D M [H]
 $>$ It is definitely other-worldly.
3. ∇ Now, this is all fantasy, but something is being trapped, this is the trap and something is emitting from it as smoke, yet what's in the center doesn't seem trapped but free, funny . . . this is all fantasy.
DW Fm, (KF?) Abst. 0—
4. ∇ Without these, the inner workings of a flower.
W CF P1
1. Well, the color mainly, I just thought of that
2. Weird men, see here's the knife in each hand.
3. This is the trap (achromatic D which had been Mars men reversed). This pink (large center Ds) could be smoke but, it's the part that would bend out toward the left and down trapping these outer forms (meaning from orange D through green D, ∇ (?) What is the trap, exactly?) Well, it is fantasy, not a plant, it's more technical . . . smoke from a building; no, I wonder . . . the colors don't suggest reality that could be described in logical terms. All this center stuff (meaning green D down through wishbone area) seems somehow free to me, funny, you'd think that was what would be in the trap, but it really seems free somehow, this outer stuff will be trapped, here on either side, this would move down catching these outer in the trap on either side.
4. The colors, but sort of shapes too, color mainly.

the cards were made). He seeks to reassure himself and the examiner that this new experience doesn't disturb him, yet he is very cautious in seeking inner cues. He describes what goes on mentally, saying in effect, how shall I take the plunge, if at all? He sticks his toes in the unknown, and wonders how shall he proceed, piecemeal or in toto. This care tells us that he is concerned, and his time reaction underscores it.

The first image is anything but a conventional one. We see now that his trepidation had at least in part to do with what kind of impression he would make, as much as the unknown qualities of the stimulus. He is concerned with the outer and the other, and this ego concern projects itself in doing something unusual. He is even

critical of his image. We can observe here a nice contrast on the vertical plane. In the middle of the remarks about his conscious processes ("the above", see Fig. 1), something emerges, the sea horse image. The appearance of this image seems to have little to do with what had been happening consciously. He immediately uses critical awareness to make a judgment about what appears (parts of the sea horses are absent). Above and below processes on the vertical plane are connected. The image is of a rather exotic form, unusual for this card, and one which has always captured man's imagination — the sea horse. This is an above-below image, a combination of the "upper-outer" land and the "down-inside" sea. Yet, as an expression of instinct, the main

feature is simulation of a higher form by a lower form, the sea horse cannot live on land. This image carries the seed of possible confusion of opposites (male-female), since it is of a species in which the male produces the offspring. Then, as if to bolster his masculine self esteem, his guard comes up with a shield or crest, and he reverses the card, quite with relief, if his comment is accepted at face value. That which was below (prior to card reversal) involves domesticated instinctive creatures (but heads only), albeit with certain odd emphasis on detail. So far, the contrasts within his instinctive world are not sharp. There is the rather weak, undeveloped, but exotic sea creature on the one hand, and the overly controlled (heads only) domesticated creatures on the other. He has a considerable problem with his passivity, about which he is guarded, and the existence of which he would deny. The cow with a crown represents the lifting of the passive, feminine and glorifying it, yet the only features of the physical sow that are present are the horns. The image implies a rather intense reaction-formation against the hostility associated with his passivity. Certain qualities which have appeared raise the question of a possible homosexual problem in the subject.

With the challenge of color (Card II), he remains superficially concerned with the horizontal plane, with the other's opinion, and he tests for what limits he might go toward, whether he follows the accepted pattern (about which he is unlikely to care very much, suggesting a psychopathic flavor). The impact of the card brings up the vaguely physiological; the moment he turns the card into the horizontal plane he is more comfortable. Yet this plane facilitates an unadaptive image, sputnik and its rocket blast. He is clearly pleased with the power implicit in the symbol and fascinated with the aggressive impulse. But there is nothing present to show any connection with the other

in this image. It is a movement away, and into outer space. He experiences power or aggression as an alien force which he is unable to direct. The instinctual world offered by the popular animals on this card, and his own association of anatomy, together with the challenge of color, all make the subject uncomfortable. He disassociates himself from involvement on any of the planes of the axes. This suggests that certain areas within his psychic world can become split off and experienced as alien. Distance between the ego and inner forces may create the experience of distance between the subject and others.

Both negative and aggressive qualities appear in association with the human figures of Card III. This popular image could allow for an easy and conventional connection with the horizontal. There is a connection, but with strongly negative overtones, and they are associated with the feminine. In addition, the quality of his logic diminishes appreciably in the up-down axis (center red), and his bright intellect ignores the result. The illogical quality here is that the center red is a bloody object because they are cannibals and they are cannibals because it is a bloody object. In the midst of a concept connected with the outer-other (popular figures), an intrusion from the below (bloody object) appears, and critical judgment (the above) evaporates quite suddenly. We have our first view of the activated archaic or autistic, and possible schizophrenic involvement. Further, there are a number of clues here about why he has a problem with his passivity (irrational orality, simulated breasts, etc.).

One of the most fascinating aspects of the record includes what the subject does with shading (see both Cards IV & VI). What he does, is perhaps the most ominous feature of the entire record, something quite bizarre, but not in the sense of the usual or common Rorschach signs of

schizophrenia. It comes out only in careful inquiry. That such is the case (not apparent in the Rorschach performance), indicates a fairly good surface adaptiveness along the horizontal plane, but looking beneath or pressing his facade, the story becomes a very different matter. This shows the sharp contrast between outer and inner (as we have seen in Baynes' description), constellated in a symbol (animal skin) that has specifically to do with adaptiveness on the horizontal. Miale (1959) has delineated the stimulus qualities of shading (such as Card IV offers) as that "which tends to evoke reactions reflecting a subject's response to nuances in experience, particularly to subtleties in relationship." Viewing the subject's responses to Cards IV and VI in this fashion, we have valuable information about his current relation to the horizontal plane. That which looks conventional, ordinary, and adaptive is found to be virtually in a state of crisis the moment one probes below the outer appearances. The adaptive qualities of the textural contact are literally reversed — he rejects what could be his greatest asset. We see his Achilles heel, he is terribly open and vulnerable in the adaptive sphere. The potential for adapted contact with others, the horizontal plane, is present, but, despite appearances, not accessible, because he feels so vulnerable. This finding alone points strongly to the losing battle of the ego forces and the vulnerability to the archaic (he closes the door to the horizontal plane on Card IV and opens it to the below aspect of the vertical on Card VI).

Even the opportunity for relief (Card V) fails to bring the subject fully back to the conventional. He sees an insect, no particular species, and this helplessness is underscored by "just something insectual," an odd term to say the least. The vague biological comment, the unadapted qualities of his use of the blot, all point to the weakness and undeveloped nature

of his . . . that involves both axes.

If there had been any doubt, his reaction to the shading on Card VI shows clearly how negative he considers the positive. He doesn't even . . . with the popular response, probably because of the . . . the texture of the skin outside in, an extreme expression of the shift from the . . . of the card to the horizontal, he selects a stimulus area that is not precisely . . . (Card V was just something "insectual," while here, he reports "a whale's mouth"). This image, following his shading panic, suggests a feeling of being engulfed or overwhelmed. It would seem that anything that suggests being open to the other creates in the subject the sense of being open to his own passivity. This is the nucleus of his crisis in relation to the horizontal plane, a powerful force removing him away from the outer-other. This is one of the most serious, if not bizarre, features in the entire record. It may be that this kind of intense loss of openness to the other (though its meanings or impact would vary with individual dynamics) may be one of the factors that makes the archaic or the autistic attractive to the schizophrenic. The pull from below becomes favored over the pull to the outer-other. Another quality connected with his behavior on Cards IV and VI is that of his being unrelated to the masculine (beneath his facade). He is fighting for his very existence, and he is in great danger of being engulfed by that to which he is the least related — his passivity, his masculine qualities, and the below.

After some effort and considerable experimentation with card axes, he produces the reversed female figures (Card VII), an "ahal" response. In some ways, this is the best image of the entire series (because of its clarity, imaginative, and creative qualities). This subject, who is so conflicted

about his passivity and easily produces the negative female image, also experiences the positive woman. He denied texture for "black fur caps," but when the idea of texture is associated with the organ of control (head) and not the instinct (animal), he shows a better adaptiveness. This is a hopeful and a very positive feature. Perhaps one can say, that if circumstances are favorable (given a certain kind of woman), the subject may feel some freedom to express his own feminine qualities and passive yearnings which, otherwise, constitute such a threat for him. If such is the case, participation along these lines would enhance the ego forces to say nothing of his self-esteem. But there are such contrasts here in the vertical plane. Rightside up, the card is only balanced rocks, "not really balanced, even though it stays up there." A pile of tenuously balanced rocks separates the subject from what little positive or flexible adaptive energy he has. It is the symmetry that enables the rocks to maintain balance, the two sides of the blot (symbolic of the two sides within) must not be differentiated (separated), but left so, or all would cave in. Things balanced, concern with balance and position are often part of the "language" of the borderline case during Rorschach. There are other points to consider here with Card VII. When the blot is rightside up, the lower D with its strongly sexual female area is prominent, to reverse the card plane may be tantamount to displacing the instinct from below. This way, tails made heads, the subject can tolerate animation and humanization. Otherwise, his image is that of an inert mass, ready to topple, and the substance turns to stone.

With Card VIII, his only really alive animal appears, but without movement. Yet, bright color does not disrupt the subject. This is one of his assets which aids the ego forces. He can tolerate a color challenge as such rather well. He "never saw a pink dog before" though; the tolerance comes

in rendering the feeling artificial or alien to the instinctual nature. The instinctive creature is domestic (a special need for the subject), but it merely looks at itself, itself is reflected back; it is not related, there are not two enlivened sides in opposition, only one, and its image caught in reflection. The creature most friendly to man (dog) is immobile and doesn't participate along the planes. While the conventional animal is enlivened, it has such a lack of purpose or goal, and the absence of movement may symbolize the emptiness in the center sphere of the axes. One sees more and more that the subject's best qualities or potential in relation to the outer—other are brittle, suspended, or cast aside.

The preference for the horizontal plane of the card is all the subject can muster under continued color stimulation (Card IX). He himself states, "I prefer to look at all of them in a longitudinal way." This is evidence for the threat of the process within him that goes on with the vertical. Finally, this card stimulates a "below" image, something from the sea depths, a plant-animal, reminiscent of the sea-land creature quality of Card I. This image's chief claim to animation is receiving food. The inadequacy implicit in this image suggests that the subject is so undeveloped below and so dependent on the intellect for air — for life itself.

With Card X, the color is abstracted into the fantasy of childhood, nothing specific, but a vague aura of the longed-for or forgotten freedom of childhood fantasy. Away from color, and the only male figures in the record appear, but alas, they are from Mars — the two fight it out with his innovation for this response — knives. This is certainly a symbol of the unrelatedness to the masculine. However, there is life in the figures, and they are certainly not passive. Though distant, a spark of aggression exists in the face of the engulfing passivity. The subject makes this not uncom-

mon image personal with his original addition (knives), and simultaneously distant (Mars). Because this same blot area takes on a different meaning when the planes are reversed, and because of certain implications for this subject, we will look more closely at it. This image from modern mythology is reminiscent of the struggle between mythological heroes of yore. Neumann (1954, p. 154) has connected dream symbols of the hero's fight with the struggle between the ego forces and the unconscious. Whether this is a symbol of the still battling ego forces or not may be an open question, but there is certainly struggle, perhaps of heroic proportions. The struggle takes place along the horizontal plane, the two sides are in violent opposition to one another, and the focus appears in "the above" sphere. Evidence of struggle is the *sine qua non* for maintaining the borderline status. At the upward end of the vertical plane, this image has some degree of vitality and energy, but when the card is reversed, the same area becomes "the trap." This locked combat (seen in relation to Cards IV and VI) points to his great fear of relation to the masculine (within and without), and the extent of the struggle about it. Still, it remains remote as an expression of viable energy or adaptive use of aggression.

The finale, set within the rich stimulus opportunities provided by Card X, is indeed unpromising. Here, at last, is the symbolic material so often associated with the paranoid schizophrenic. But note again the "language of the borderline," and the care with which he says in effect, "I have to accept this even though I know it is bizarre." The borderline patient is able to interpose a certain narrow pattern of experience and reaction between the stimulus and the response. The experience is schizophrenic, the response to it is more appropriate and adapted. However, the lure of this image for the subject is clear

enough. He gives in to the lure by deliberately suspending the criteria of reality necessary to keep related to the outer-other. The criteria of reality are not discarded, but set aside for the duration of the subject's involvement with his image. We must take careful note of the planes of the blot, for here we have the "schizophrenic break." Let us look at the process and its conditions. Observe that the upright plane is reversed, the below is above, and note the distinction the subject himself makes between the above, the below, the outer and inner, the free and the trapped. This response offers a most graphic representation for what we have been trying to delineate—process level activity in the borderline schizophrenic. It is a mirror of what could so easily take place, but what in this subject *hasn't* taken place on a behavioral level. He is lured but he is, so far, keeping his distance. This is Rorschach evidence for what Baynes meant when he described "archaic elements which flood the mind, but which he cannot express." We see the movement up the vertical from below, the archaic material emerging and becoming "free" (freely conscious) and the movement on the horizontal becoming trapped—this is the image, the symbol-projection of the deadly quandary of the borderline schizophrenic. "The trap" is the activation of the archaic, which lures the conscious away from the authentic or outwardly oriented personality (as, when the ego battle goes up in smoke, the personality is trapped in the unconscious).

DISCUSSION

We must now look at what we have in terms of the framework we have presented. Where is the focus in this subject's struggle with schizophrenic elements? The usual terms associated with interpretation (defenses, dynamics) have been largely avoided. This has been necessary so as not to obscure the dynamic we have called "process".

The subject has not given up a number of things which connect him to the horizontal plane. He is fighting to keep his orientation toward the other. We have seen his concern with his own balance and his alertness about things that might disturb this balance. Evidence for some development in the spheres of contact and adaptive awareness (as contrasted, for example, with the instinctive sphere) is indicated. We have material reflecting fight, battle, and resistance. He makes a deliberate attempt to keep the archaic distant and alien. After grappling with it, he can return to the tangible world. By and large, conscious control makes a good showing for itself. In certain important dynamic areas, bits of positive possibilities shine through the gloom of the negative aspects.

Yet, his great vulnerability is what is happening on the vertical plane. At this point, his adaptive qualities are experienced as liability. That which would best keep him connected with the horizontal has become his most vulnerable spot (shown by his reaction to shading). The largely unformed and pervasive weakness in the instinctual world offers little support. Precarious balance is suggested at every hand. When unconscious or archaic products emerge, they have no connection with specific forms (they are just vague abstractions as shown on Cards III and X); hence, the rational is alien to them (thus, his rational need to keep *them* alien). Feeling (color) and contact (texture) brought out bizarre products, the former during the performance, and the latter during the inquiry. One would be inclined to predict that this is a subject whose ego forces are fighting a losing battle with the unconscious. His chief hope is to accept again the adaptive potential which he is now so inclined to discard (response to shading, Cards IV and VI). The focus of the process is now in the borderland; this may be replaced by the "never-never land."

Shapiro (1954) has suggested that in borderline Rorschach records, it is not so much the presence of autism, but the way of handling it that becomes crucial. He added that we need to know more about ego functions resistant to disruption. It may well be that it is not only the way of handling autistic material that becomes crucial, but also the direction and strength of forces. We saw how our subject "handled" autistic material, and it looked fairly positive, but this, framed against the rest of the picture, made the outlook dim. We can't be sure that ego functions resistant to disruption are in themselves those which block the movement of the psychotic forces. Relation between cause and effect is as yet the great unknown in schizophrenia.

In delineating Jung's theory of schizophrenia, Baynes (1940, p. 39) described two distinct parts within the schizophrenic process. He considered lowering the level of consciousness (critical judgment and conscious awareness are diminished) the first part, and activation of the archaic (autistic) contents of the unconscious as the second. What we need to know is what happens to the ego forces when the archaic is activated, what is the nature of the effect? Jung (1960, p. 233-249) looked at the development of schizophrenia on a process level. He stated in the early part of the century that the neurotic fights for supremacy of ego consciousness, whereas the psychotic identifies with the morbid elements. We have looked at our borderline case in a similar way from the standpoint of Rorschach process. In such a case, when guideposts are few and the stimuli are relatively ambiguous, conscious control appears to diminish in the face of the fascination with "the below," the archaic activation (as, in our subject, content material on Card X). In the borderline case, the degree to which this occurs determines the extent of the disturbance. It may occur only minimally or pervasively. The out-

ward behavioral manifestations may not always coincide with the degree of imbalance between the forces within the psyche. One may not be justified in calling the subject schizophrenic when the ego forces prevail. However, the very presence of the archaic, the autistic, may have a profound effect on the personality. We know little about what the effect may be. It has been suggested that a process analysis is helpful in delineating the battle and predicting the outcome.

It would seem that in the borderline case, when archaic material appears in the Rorschach record, these products are at least partially activated. That correlates on the behavioral level cannot be demonstrated does not alter the presence of this condition. In these cases, we are sometimes faced with a more subjective (Rorschach) versus a more objective (psychiatric) definition of schizophrenia. The term subjective is meant here to describe the source from which the data came (the more subjective side of the subject), and it does not refer to a quality used in making a judgment. In like manner, objective does not here mean "more scientific," but it refers to the source of the data, the subject's outer, more formalized world. Both methods include ways of looking at or inferring information about both areas of reference. Each may have its own validity and they should be interdependent. We should have no hesitation in using a borderline schizophrenia designation. When the battle is near at hand, even though one may not see behavioral correlates so clearly, the schizophrenic features in the person must be taken into account. These people have a particular kind of battle that neither the neurotic nor the schizophrenic has. In one way or another, they are vitally involved with their struggle. Attention to exactly what happens in these cases may take us a long way toward understanding schizophrenia.

Great concern with the nature of defense patterns has not brought forth

an answer. In fact, though defense patterns are designated by "scientific terms," we seldom realize how close the functions we ascribe to them resemble primitive magic (we can't be sure that such and such a defense "insulates against the onslaught of the autistic," though we sometimes speak as if we were sure). We may be too glib about their protective functions, especially when we don't know the effect the archaic has upon them. In his profound treatment of the situation surrounding the borderline schizophrenic, Baynes (1940, p. 43) has this to say:

"According to my experience, patients who eventually cross the line from neurosis to psychosis go voluntarily. Conflict is usually acute up to the point when the renegade hypothesis is more or less consciously accepted. It is rather as though an essential part of the personality had migrated to 'the other side,' and eventually the rest of the personality is bound to follow, and usually without great protest."

From the considerations in this paper, it should be clear that prediction, diagnosis, and accuracy in the borderline case are by no means a simple matter. The burden of evidence should lie on what has been found in the psyche of the subject concerned and not in the label applied. We run the risk of being superficial or hasty in our pressure to "solve" a diagnostic problem. At least in the borderline case, diagnosis based exclusively on the presence of autistic material in the Rorschach, or exclusively on what can be demonstrated behaviorally, may both miss the mark.

Perhaps it can be seen from the points considered herein how validity studies about the Rorschach sometime fail to take vital points into consideration. Psychiatric diagnosis is not always the appropriate standard of comparison or criterion for predictions based on Rorschach findings. The presence or absence of certain commonly expected "signs of schizophrenia" (like minus form and confabulations) may not necessarily be the only deciding factors in Ror-

schach diagnosis. If we do not make the appropriate distinctions about the sources for our data, how can we expect to design adequate validity studies? If our forms of investigation of a disorder can become more or less split off from one another in emphasis (objective or behavioral side of the subject versus the inner, subjective side), how can we talk accurately about the split within a subject's psyche? Behavior within is just as important data about a particular individual as data in relation to behavior without. The inner and outer worlds of an individual are always intimately connected, even though the connection may not always be obvious or tangible. The approach to the inner world represents a thorny problem. There is an enormous difference between "subjectivity" and critical judgment about the subjective. Fruitful research in the realm of psychic disorder should include ways of bringing the two together; the one should not be neglected in favor of the other.

SUMMARY

Problems centering around communication of Rorschach findings in borderline schizophrenic cases have been considered. Attention was directed to the not infrequent disparity between absence of clinical symptoms as reported from psychiatric examination, and the presence of schizophrenic material as reported from Rorschach ex-

amination. The Rorschach record of a borderline schizophrenic case was presented and analyzed. This interpretation was used to illustrate and define an approach to the Rorschach material which the writer called "process analysis." A theoretical schema was proposed to enable the examiner to grasp the direction of movement in what was described as two planes of struggle characteristic of the borderline case. The findings were related to certain theoretical lines associated with schizophrenia. Questions were raised about some of the prevailing concepts in relation to this disorder. Attention was directed toward sharpening diagnostic procedures in the borderline case.

REFERENCES

- Baynes, H. G. *Mythology of the soul*. Baltimore: Williams & Wilkins, 1940.
- Bleuler, E. *Affektivitat, suggestibilitat, paranoia*. Halle: Marhold, 1926.
- Forer, B. R. The latency of latent schizophrenia. *J. proj. Tech.*, 1950, 14, 297-302.
- Jung, C. G. *The psychogenesis of mental disease*. New York: Pantheon, 1960 (1939).
- Miale, Florence. An approach to projective material. Doctoral dissertation, The New School for Social Research, New York, 1959.
- Neumann, E. *The origins and history of consciousness*. New York: Pantheon, 1954.
- Schafer, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- Shapiro, D. Special problems of testing borderline psychotics. *J. proj. Tech.*, 1954, 18, 387-394.

Received March 12, 1962

The Performance of "Chronic Schizophrenic" Patients on Piotrowski's Rorschach Sign List for Organic CNS Pathology¹

STEPHEN NEUFELT, ALAN G. NEWMAN

Toronto Psychiatric Hospital

AND DOUGLAS A. QUINCY

Ontario Hospital, Toronto

Piotrowski's (1957) sign list for the diagnosis of organic CNS pathology purports to classify patients into two categories, namely, those with, and those without brain damage. This sign list has been both confirmed (Bech, 1946; Fisher, 1955) and criticized (Tatney, 1944; Kisker, 1944; Hughes, 1950; Hertz and Loehrke, 1954), so that its status as a useful means of prediction remains in doubt. This, however, does not mean that there is clear reason to doubt the validity of the sign list. Nearly all the studies using the Piotrowski list show that high scores on this sign list tend to occur less frequently in nonorganics than they do in organics. However, criticism of the scale suggests that it tends to identify many organics as nonorganic (false negatives).

PROBLEM

The literature points to five general problems which require attention before adequate information regarding the usefulness of the Piotrowski list can be obtained. These problems concern: (a) the selection of the criterion ("organic") group, (b) the selection of control groups, (c) the scoring of the sign list, (d) the selection of the cut-off point, and (e) the discriminative ability of the separate items of the

sign list. These problems are discussed below.

The criterion group has frequently proved problematic in studies of the Piotrowski list. Hertz, Jones, and Kisker (1944), working with lobotomized patients, appear to have found relatively low frequency of Piotrowski signs in their "organic" group. Nadel (1938), however, seemed to find that frontal lesions were more likely to lead to positive Piotrowski tests than lesions elsewhere in the brain. Hartman and Erickson (1940) report that the location of tumour has relatively insignificant effects upon the Rorschach performance. This confusion would appear to be supported by Birch and Diller's (1959) observation that a relatively high relationship exists between the gross number of medical signs found in neurological examination and the number of Piotrowski signs in the Rorschach. It must at least be stated that it is not yet known to what extent and in what ways (if any) the Piotrowski score is affected by size, location, or source of brain damage. Consequently, for the present, it would seem most practical, in studies concerned with the Piotrowski list, to select criterion groups in such a way that it may be assumed that the type of damage present will be nonspecific as to location and etiology in case the type and/or location of damage is relevant to the Piotrowski score.

The control groups which are used in studies of the Piotrowski list are, if anything, more critical to the outcomes of the studies than the criterion groups selected. Fisher, et al (1955) report that the number of nonorganics

¹This study was supported by funds provided under the Federal-Provincial Mental Health Grant, No. 605-5-277, of the Canadian National Health Grants programme. Grateful acknowledgment is also extended to Dr. A. B. Stokes, Head, Department of Psychiatry, Univer. of Toronto for his support, and to the diverse facilities provided by the hospitals, clinics and agencies of the Ontario Department of Health and by many other community and business organizations.

scoring as "organic" (false positives) in their sample was relatively small. Indeed, they show that in their sample the Piotrowski list performed an overall discrimination of "organics" and "nonorganics" which was approximately as efficient as the best neurological test (the EEG) they employed, due to the tendency of the Piotrowski scale to yield relatively low numbers of false positives, and in spite of its tendency to yield relatively high numbers of false negatives. Tarcsay (1944), however, found that a relatively large number of her nonorganic (psychotic) subjects scored above the cut-off point on the "Piotrowski list" (which she altered and scored in an extremely liberal manner), a fact from which she concluded that rather than estimating organic impairment, the Piotrowski list yields a score relevant to psychosis in general. In this connection, it is worthy of note that Bash (1946) cast doubt upon Tarcsay's conclusions, apparently to the satisfaction of the Central European Rorschach school (Bohm, 1957), by showing that among other groups, a large group of his neurotics (73%) also scored high on the "Piotrowski list" (as scored according to Tarcsay's instructions), thus demonstrating that the sign list, as Tarcsay used it, was not specific for psychosis. But it would also seem that Bash's results support the conclusion that the list, as changed by Tarcsay, is not specific for brain damage either. However, it seems to the present authors that the large number of false positives reported by Tarcsay and Bash is due at least in part to the changes introduced by Tarcsay into the scoring of the list, and that their results are therefore not strictly comparable to the results reported by other authors of studies using the Piotrowski list.

Piotrowski (1937) and Fisher, et al (1955) both assumed that the critical contrast for the scale should be between true "organics" and nonorganic neurotics whose symptoms simulate those of the organic. Since this

type of neurotic posed a serious problem for neurological investigation, it seemed to Piotrowski and to Fisher that this discrimination would be a useful one to achieve. Many other authors have used normal subjects as controls, but these were apparently used mainly to show group differences, rather than to estimate the tendency of the scale to show *individual* false positives. However, there is some reason to suspect that the fact that Piotrowski failed to direct sufficient attention to psychotics in the selection of his control group, and instead focused on neurotics, may have led to an error in assigning an identity to the function which the scale estimates.

Birch and Diller (1959) state that the Rorschach "does not appear to be sensitive to brain damage *per se*, but it does reveal 'organicity', one of the behavioural consequences of brain damage. We would prefer to think of 'organicity' primarily as an impairment in organization of perceptions". Tarcsay's (1944) findings, in spite of the limitations arising from her re-interpretation of the Piotrowski signs, might still suggest the possibility that "organicity" may also be a behavioural consequence of psychosis.

Further support for this more general way of viewing the factor estimated by the Piotrowski sign list has emerged from our own investigations. During a series of studies concerned with the derivation of Rorschach scales to measure attributes of chronic psychosis, the present authors observed that the Piotrowski signs appeared quite frequently among the signs of "regression" accompanying chronicity in psychosis (see Neiger, et al, 1961b). Although other more efficient signs were found for this purpose, the fact that some success was attained using the Piotrowski list in the prediction of chronic schizophrenia², suggested that an attempt should be made to re-

²That is, schizophrenics tended quite frequently to score as "organic" (more than

assess that scale's tendency to make false positive errors of prediction. The need to investigate the misplacement rate of the scale in schizophrenia seemed particularly great in view of the increased risk in making errors in prediction resulting from the relatively large size of this group of patients within the clinical population to which the Rorschach is likely to be applied.

The scoring of the sign list is unquestionably a very important problem which nevertheless seldom receives much comment in the literature. Fisher, et al, (1955) used protocols administered and scored by psychologists at various levels of skill and experience with the Rorschach. Fisher and his colleagues rescored the protocols on the assumption that the original scorers might not have scored the protocols in a consistent manner.

It has been our observation that the Piotrowski signs are somewhat ambiguously defined, so that we would assume that not all authors understand the scores in the same way. Certainly, some authors have explicitly used the sign list in a different way from that laid down by Piotrowski. Tarcsay (1944) so radically altered the sign list's scoring criteria that many normal subjects would probably obtain high scores on the list. But more subtle scoring differences seem almost certain to creep in to the scoring practices of various workers. The number of responses (sign #1:R) the number of human movement responses (sign #3), the presence or absence of colour naming (sign #4), the scoring of form level (sign #5), and the number of popular responses (sign #6), will all depend to some extent upon the widely varying definitions of these responses used in the scoring system that happens to be used by a given worker. The time per response (sign #2: T/R) will vary according to the speed of writing of the administrator and the passivity required by the sys-

tem of administration he uses. Finally, the scoring criteria for Perplexity (sign #8: PNC), Impotence (sign #9: Imp), Automatic Phrases (sign #10: Aut-Phr), and to some extent, Repetition (sign #7: Rpt) are not very clearly outlined by Piotrowski, and are understood slightly differently by various authors. Our own scoring of the Piotrowski list will tend to show longer response times (more positives on the sign), fewer human movement responses (more positives on the sign), and more popular responses (fewer positives) than the average scoring by persons using other systems of Rorschach scoring.

The selection of the cut-off point cannot be altered if any semblance of comparability is to be assumed from study to study. However, a few observations are pertinent to this aspect of the scale. In the first place, it is clear that the appropriate cut-off point to perform the most efficient group discrimination will vary depending upon the conservatism or liberalism of the scoring practices used. And, in addition, it is obvious that the farther the cut-off point is set *up* on the scale the fewer will be the scale's false positives, and the farther *down* on the scale it is placed the fewer will be the scale's false negatives. Present indications, based on Fisher's (1955) observations, would suggest that the cut-off point is rather high, and that consequently the scale yields relatively few false positives. However, as Hunt (1940) has indicated, it may also yield relatively few correct positives. The problem appears to be that the range of uncertainty in diagnosis is fairly large, and that the cut-off point cannot be elevated greatly due to the ceiling effect created by the relative infrequency of occurrence of most of the signs. The latter problem cannot be solved readily, but the former problem might be dealt with to some extent by an item analysis of the signs since large ranges of uncertainty are usually due to measurement error

which can result from inclusion of poorly discriminating items.

The relevance of the items of the Piotrowski list to the criterion prediction has frequently been subjected to rough analysis, usually by inspection. Most authors who comment on the signs of the Piotrowski list point to Impotence and Perplexity as the main signs which are specific for brain damage. But the fact remains that very little direct information is available concerning the performances of various groups of subjects on the specific Piotrowski signs, although specific signs contained in the Piotrowski list have been reported in syndromes other than that of brain damage.

The present authors have reported elsewhere (Neiger, et al, 1961b, 1961c) upon the behaviour of the total number of responses, the form level, the number of sharply perceived human movement responses, colour naming, and of popular responses in schizophrenic illnesses. They have also observed that the entire Piotrowski sign list could be used to yield predictions of chronicity in psychosis, although some of the signs did not contribute much to the efficiency of the predictions. It would therefore seem desirable to attempt to supply some further information concerning the separate signs of the list in order to ascertain which of the signs are specifically useful for the criterion prediction of brain damage per se.

It is therefore proposed to investigate the distributions of Piotrowski scores and the occurrences of the separate Piotrowski signs in groups of normal, chronic schizophrenic, and "organic" subjects.

METHOD

Subjects

A total of 140 subjects were used in this investigation. All subjects were between the ages of 20 and 40 years, and had been educated since Grade 1 in an English-speaking country.

Four groups of normal subjects were assembled from our files of research

protocols. One group of 20 female subjects of varied educational levels was used, and three groups each containing 20 male subjects were selected. The three male groups were selected according to their educational attainments, in order to observe the effects, if any, of education upon the Piotrowski signs. One group of male subjects had attained a university level of education ("high" normals), one group had attained an educational level between grades 10 and 12 ("medium" normals), and one group had attained an educational level of grade 8 or 9 ("low" normals). "Normal" here means a person who was approached to serve as a subject for testing, rather than having been tested under conditions of seeking any kind of help. The normal protocols were obtained from several large industries which made their employees available for research testing, and from Rorschach students who tested their acquaintances during practice testing. There was no attempt to screen the normal subjects for psychopathology.

Two groups of chronic schizophrenic patients were assembled as follows: 20 female chronic schizophrenic patients of varied educational level; and 20 male chronic schizophrenic patients of varied educational level. Of the male chronic schizophrenics, 10 were inactive ward patients ("nonworking" schizophrenics) and 10 were ward patients who were employed usefully in performing various simple duties around the hospital ("working" schizophrenics). The schizophrenic patients were tested in several Ontario Hospitals. All the patients were diagnosed independently by psychiatrists without disagreement. None of the schizophrenic patients had undergone lobotomy, and all were screened clinically to rule out suspicion of organic involvement. No patient who was receiving drug treatment was tested unless the original symptoms were still apparent, and patients were not tested until at least two weeks following ECT or ICT, if any. All the schizo-

patients had been hospitalized for at least two years, and in most cases for more than two years.

One group of "organic" subjects was assembled as follows: 20 male chronic idiopathic epileptic patients, of whom 10 exhibited no clinical features of psychosis (Internat. classif.: 353.1), and 10 exhibited psychotic features secondary to the epilepsy (308.1). All the epileptic patients had been in hospital for many years, unable to care for themselves as a result of their illness. It was assumed that, regardless of the cause(s) of idiopathic epilepsy, epileptic discharges frequently repeated over many years should eventually cause some disturbance in the functional harmony of the brain as a whole and between its several regulating systems, the effects of which would be the same upon behaviour as the effects of brain damage. Such disturbances should be nonspecific as to location and extent of damage.

Hypotheses

In the light of the findings we have reported elsewhere (Neiger, et al, 1961b, 1961c) we would expect that on the total Piotrowski scale the organic patients would tend to score highest, the schizophrenics would tend to score only slightly lower, and the normals would tend to score significantly below both the other groups.

We would further expect, in the light of our findings referred to above, that the various subgroups would show increasing central score tendencies on the Piotrowski list in the following order: (a) high level normals, (b) medium level normals, (c) low level normals (because of expected educational effects); (d) "working" chronic schizophrenics, (e) "non-working" chronic schizophrenics (on the assumption that the former group is less incapacitated than the latter); (f) nonpsychotic deteriorated epileptics, and finally (g) psychotic deteriorated epileptics (on the assumption of additive effects of incapacitation

due to psychosis and organic deficit).

With regard to the individual items of the scale, we would expect that there would be no differences between the "organics" and the schizophrenics on the low number of responses, the low number of human movement responses, the presence of colour naming, the low form plus percent, and the low percentage of popular responses, whereas the "organics" would tend to show more Impotence, Perplexity, and Repetition reactions than the chronic schizophrenics.

Procedure

The Rorschachs were administered individually according to the procedure recommended for the original continental method of Rorschach analysis (Bohm, 1957; Neiger, et al, 1961a). The records were scored for the Piotrowski signs according to our best understanding of Piotrowski's scoring instructions, and the scoring was checked by another investigator.³

The groups were compared by the chi-square method for the numbers of subjects misplaced, using a cut-off point of 5 or more signs. The correlation between predicted and observed rank orders of the subgroups were tested with the Spearman rank correlation method, using the median score for each group. The groups were also compared for each item on the sign list.

RESULTS

Table I presents the main findings of this study. It shows the number of subjects in each of the groups who obtained scores of 4 or less, 5 or more, and 6 or more⁴ on the scale.

Table II presents the results of the comparison between groups using a cut-off point of 5 or more. It will be observed that significant values of chi-square were obtained in the comparisons of normals and epileptics, and

³However, note our remarks under Problem, concerning probable sources of difference.

⁴This cut-off point is suggested as a supplementary one by Piotrowski's (1937) data.

TABLE I—Number of Subjects in Each of the Groups Studied Having Piotrowski Scores of 4 or Less, 5 or More, and 6 or More

Group	N	Number of Ss with scores of:			% Misplaced 5%
		4 or less	5 or more	6 or more	
Normal Females	20	19	1	0	
Normal Males:					
University Educ.	20	20	0	0	0%
Grades 10-12 Educ.	20	20	0	0	0%
Grades 8-9 Educ.	20	19	1	0	5%
Chronic Schizophrenics:					
Females	20	11	9	2	45%*
Males:					
"working"	10	7	3	1	30%*
"nonworking"	10	4	6	3	60%*
Epileptics:					
Males:					
nonpsychotic	10	4	6	4	40%
psychotic	10	1	9	5	10%

* "Misplaced" here means the percentage of chronic schizophrenics scoring in the "organic" range of the scale, i.e., having 5 or more Piotrowski signs.

TABLE II—Tests of Significance of Between Groups Differences: Chi-square Tests Using the Cut-off score of 5 or more signs present

Groups Compared	X ²	p
High vs Medium vs Low Education Normals	1.03	ns
High Educ. Normals vs Chronic Schizophrenics	11.6	.001
Medium Educ. Normals vs Chronic Schizophrenics	11.6	.001
Low Educ. Normals vs Chronic Schizophrenics	8.54	.01
High Educ. Normals vs Epileptics	24.0	.001
Medium Educ. Normals vs Epileptics	24.0	.001
Low Educ. Normals vs Epileptics	16.25	.001
Normals vs Chronic Schizophrenics: Female	8.53	.01
Chronic Schizophrenics vs Epileptics	3.75	ns
"Working" vs "Nonworking" Schizophrenics	1.82	ns
Psychotic vs Nonpsychotic Epileptics	2.40	ns

of normals and schizophrenics. The scores obtained by the various normal groups do not differ significantly from each other. Nor do the scores obtained by the epileptics and the schizophrenics differ significantly from one another. However, the Mann-Whitney *U* test, which employs more of the distribution data present in the material than does the chi-square test, yields a significant value of *U* ($U=129$, $p=.026$) between the epi-

leptics and the chronic schizophrenic patients. Table II also indicates that the misplacement frequencies do not differ significantly for (a) the three educational levels within the normal male group, (b) the "working" and the "nonworking" schizophrenic subjects, and (c) the psychotic epileptics and the nonpsychotic epileptics. Mann-Whitney *U* tests applied to the last two group comparisons also fail to attain significance ("working" vs. "nonworking" schizophrenics: $U=29.5$, $p=.053$; "psychotic" vs. "nonpsychotic" epileptics: $U=37.5$, $p=.06$). However, the relationship between the expected and the observed rank orders of the groups shows no reversals ($r_s=1.00$), and is therefore significant.

Table III presents the scattergram of scores for each group. It will be seen that the epileptics appear to occupy a slightly higher position on the scale than the chronic schizophrenics, and that there is a slow but consistent increase in the medians for the various groups when they are listed in order according to the positions they would be expected to assume under the hypothesis of the investigation.

Table IV tabulates the frequencies with which the specific Piotrowski items appeared in each of the groups. It will be observed that only three of the signs (*R*, *T/R*, *Rpt*) show signifi-

TABLE III. Frequency Distributions of Scores Obtained on the Piotrowski List in the Three Types of Subjects Used in This Study

Groups	Piotrowski Score								N	Mdn
	0	1	2	3	4	5	6	7		
Normals										
Female	1	7	10	1		1			20	1.7
Male:										
High Education	2	4	12	2					20	1.8
Med. Education		6	5	5	4				20	2.3
Low Education	3	5	2	5	4	1			20	2.5
Chronic Schizophrenics:										
Female		1	4	3	3	7	1	1	20	4.2
Male:										
"Working"			2	1	4	2	1		10	4.0
"Nonworking"				1	3	3	2	1	10	4.8
Epileptics:										
Male:										
Nonpsychotic				2	2	2	3	1	10	5.0
Psychotic					1	4	4	1	10	5.5

cant differences in frequency of occurrence between the chronic schizophrenics and the epileptics, and that one of these (R) reverses the direction of the difference from the one expected under the hypothesis that the Piotrowski list estimates organic impairment.

DISCUSSION

Table I shows that the Piotrowski list is effective in (a) minimizing false positives among the normal groups used: only 2 subjects out of 80 obtained scores of greater than 4 on the list, and (b) keeping to a fairly low proportion the number of epileptics identified as nonorganic by the scale. But it also shows that *about half of the chronic schizophrenics* were identified as organic using a cut-off score of five or more. Although this last figure may be reduced to about 10% if the cut-off point is placed at six, this procedure would also serve to "screen out" over half of the epileptics in our data. It would appear, then, that within the limits imposed by the subjects used in this study, the hypothesis under investigation has been confirmed. That is, it seems probable that chronic schizophrenic patients tend to fall at positions on the Piotrowski scale which are fairly close to the positions occupied by organic patients. This will be seen by a glance at Table III which presented the frequency dis-

tributions for the groups used in this study.

But other observations can also be made on the basis of Table III. It will be observed that there appears to be a slight but consistent and significant trend for (a) the lower education normals to obtain higher scores than the higher education normals, (b) the "nonworking" schizophrenic patients to obtain higher scores than the "working" schizophrenic patients, and (c) the psychotic epileptics to obtain higher scores than the nonpsychotic epileptics. It might be argued that these contrasted groups differ from one another in terms of other psychologically relevant variables. (a) Persons with high education as opposed to persons with low education are more likely to have good abstraction ability, to have developed considerable cognitive complexity, to have developed some degree of independence of thought and action, and to have good intelligence. (b) In view of the characteristics by means of which they were distinguished initially, "working" as opposed to "nonworking" chronic schizophrenic patients are more likely to be able somewhat to control their behaviour, to be somewhat reliable and persistent, to be able to function somewhat independently, to be able to use some low level skills effectively, and to have retained some

measure of cognitive organization and complexity. (c) The concept of psychosis added to the diagnostic label would suggest that "nonpsychotic" epileptics as compared with "psychotic" epileptics are more likely to have retained some cognitive organization, some more or less adequate object relations, and some semblance of perceptual and conceptual differentiation. It appears to us that, in a very general sense, it is these attributes which are lost first in deteriorative or regressive psychiatric illness. *We have therefore assumed that the concept which should be used in order to identify the factor estimated by the Piotrowski scale is that of "regression".*

Table IV supplies further data in support of the view that some of the Piotrowski items are not specific to brain damage, but are more general

signs of psychotic level disturbances in functioning. From this listing of the frequencies of occurrence of each sign for every group, it will be seen that one of the signs (R) appears to be more specific to the chronic schizophrenics than to the epileptics, three signs (Rpt, Plx, T/R) appear to be more specific for epileptics than for chronic schizophrenics, two signs (Pop%, Imp) appear to be specific to no group in particular, and the rest (Human Movement, Colour Naming, F+%, Automatic Phrases) appear to occur as frequently in epileptics as in chronic schizophrenics although they occur less frequently in normals. These observations are of interest particularly in view of our findings reported elsewhere (Neiger, et al 1961b) concerning Rorschach factors in chronic psychosis. In this other work we found that chronic psychosis could be

TABLE IV—Item Analysis: Frequencies of Occurrence of Each Sign in the Various Groups Investigated in the Present Study, and Chi-Square Tests of Significance of Differences between Frequencies of Occurrence in the Groups of Male Chronic Schizophrenics and Epileptics

Groups	Piotrowski Signs ^a									
Normal	I	II	III	IV	V	VI	VII	VIII	IX	X
Female										
Male:	1	14	9	0	1	9	1	0	0	0
High Education	0	9	9	0	1	13	0	0	0	1
Med. Education	2	19	10	1	0	12	1	2	0	0
Low Education	5 ^b	5	14 ^b	0	4 ^b	7	4 ^b	5 ^b	0	0
Abnormal										
Female Chron. Schizophr.	14	11	19	1	11	15	2	2	1	2
Male, Chronic Schizophr.:										
"Working"	7	8	8	1	7	4	1	2	0	1
"Nonworking"	9	6	9	3	8	8	3	2	0	1
Male, Epileptic:										
Nonpsychotic	3	9	9	2	7	7	4	6	0	2
Psychotic	5	10	9	2	9	8	6	3	1	1
Chron. Schizophr. vs Epileptic (males)										
X ² Tests	6.66** 4.34**					3.95**				

^a I: Number of Responses ≤ 15 ; II: Time per Response $\geq 60''$; III: Human Movement $\leq 1+$; IV: Colour Naming present; V: F+% $\leq 70\%$; VI: Popular % $\leq 25\%$; VII: Repetition present; VIII: Perplexity present; IX: Impotence present; X: Automatic Phrases present.

^b It will be observed that some of the signs (I, III, V, VII, VIII) appear to be related to the educational level of the normal subjects. For a discussion of education effects on some Rorschach scores, see Neiger, et al, 1961b.

^c Direction is different from that expected: the chronic schizophrenic group contains more cases involving low R than the epileptic group.

* Significant at .01 level.

** Significant at .05 level.

predicted reliably using only R, F+%, and Colour Naming for male subjects, and using these three signs along with corrections for three other signs (including Human Movement) for female subjects. Since, in our other work, we have identified these findings with the concept of regression of the type occurring in chronic psychosis, it will be clear that our attitudes are committed to the view that some of the variance in the Piotrowski scale is a function of a factor other than the specific factor of brain damage.⁵ That is, it would seem to us that if brain damage *per se* is to be predicted, it would be desirable to avoid the pitfalls of including identifiable error in the measurement by including items which are not specific to brain damaged patients.

SUMMARY

A group of 80 normal subjects, containing subgroups with specified levels of educational achievement, a group of 40 chronic schizophrenic patients, and a group of 20 idiopathic epileptics were compared for their performances on Piotrowski's Rorschach sign list for the diagnosis of organic CNS pathology.

It was observed that, although the chronic schizophrenics tended to score slightly lower on the scale than the epileptics, about half of the schizophrenics obtained scores in the "organic" range. This finding tends to contradict previous findings which reported a low false positive rate for the Piotrowski sign list. At the same time, this finding supports the hypothesis, derived from other work by the present authors and from previous investigations with this sign list, that the Piotrowski list yields an estimate of the regressive process in chronic psychotic illness.

⁵Of course, as an alternative viewpoint it is sometimes asserted that at least some forms of schizophrenia may have organic roots. The possible validity of this view is not prejudicial to the present argument if the Piotrowski scale is to be used to distinguish currently recognized diagnostic groups.

Examination of the discriminating ability of the individual signs of the Piotrowski list indicates that only a relatively few signs supply significant group discriminations when the control group is comprised of chronic schizophrenics. Considerable caution is indicated in the application of this scale to a psychiatric population, if it is to be used to predict brain damage.

REFERENCES

- Bash, K. W. Ueber den differentialdiagnostischen Wert der Piotrowski-Zeichen und anderer Zeichen im Rorschach-Versuch. *Experimentia*, 1946, 2.
- Birch, H. G. & Diller, L. Rorschach signs of organicity: a physiological basis for perceptual disturbances. *J. proj. Tech.*, 1959, 23, 184-197.
- Bohm, E. *Lehrbuch der Rorschach Psychodiagnostik*. Bern: Huber, 1957; Abbr. Engl. Edit., New York: Grune & Stratton, 1958.
- Fisher, J., Gonda, T. A., & Little, K. B. The Rorschach and central nervous system pathology: a cross-validation study. *Amer. J. Psychiat.*, 1955, 111, 487-492.
- Harrower-Erickson, Molly R. Personality changes accompanying cerebral lesions: I Rorschach studies on patients with cerebral tumors. *Arch. Neur. Psychiat.*, 1940, 43, 859-890.
- Hertz, Marguerite R., & Lochrke, Leah M. The application of the Piotrowski and the Hughes signs of organic defect to a group of patients suffering from post-traumatic encephalopathy. *J. proj. Tech.*, 1954, 18, 183-196.
- Hughes, R. M. A factor analysis of Rorschach diagnostic signs. *J. gen. Psychol.*, 1950, 43, 85-103.
- Hunt, T. The application of the Rorschach test and a word association test to patients undergoing prefrontal lobotomy. *Psychol. Bull.*, 1940, 37, 546. (abstr.)
- Kisker, G. W. The Rorschach analysis of psychotics subjected to neurosurgical interruption of the thalamo-cortical projections. *Psychiat. Quart.*, 1944, 18, 43-52.
- Nadel, A. B. A qualitative analysis of behaviour following cerebral lesions, diagnosed as primarily affecting the frontal lobes. *Arch. Neurol. N. Y.*, 1938, 32, No. 224.
- Neiger, S., Quirk, D. A. & Slemmon, A. G. *Manual of the Rorschach psychodiagnostic*. Toronto, author, 1961.
- Neiger, S., Slemmon, A. G. & Quirk, D. A. Scales of psychotic regression from the Rorschach test. Toronto, author, 1961.
- Neiger, S., Slemmon, A. G. & Quirk, D. A. Rorschach indices of common associations from popular responses: Part III. Valid-

tion and cross validation. Toronto, author, 1961.

Piotrowski, Z. A. The Rorschach inkblot method in organic disturbances of the central nervous system. *J. nerv. ment. Dis.*, 1937, 86, 525-537.

Tarcsay, Isabella. *Grundriss der Psychodiagnostik*. Zurich: Rascher, 1944.

Received December 6, 1961

Revision received June 18, 1962

Base Rates in the Prediction of Suicide: A Note on Appelbaum's and Holzman's "The Color-Shading Response and Suicide"

JEROME D. PAUKER
University of Minnesota, School of Medicine

Appelbaum and Holzman (1962) present the following results (part of their Table I) in reporting on the predictive and postdictive validity of a suggested Rorschach sign (color-shading) of suicide:

TABLE I—Incidence of the Color-Shading Determinant In Suicidal and Non-Suicidal Subjects

		Color-Shading	
		Present	Absent
Experimental Groups	Suicides	27	3
	Attempted Suicides	32	7
	Sum	59	10
Control Group	Non-Suicidal Patients	18	78

The base rate for suicide (actual plus attempted) for psychiatric patients in general at the C. F. Menninger Memorial Hospital was found to be 28%. Appelbaum and Holzman interpret their results as follows (1962, p. 158):

Table I indicates that the test sign correctly identifies eighty-six per cent of suicidal patients, a significant increase in predictive power beyond the twenty-eight per cent base rate.

The wrong figures are used, however, to compare with the base rate. Since the test sign is used to pick out

suicides from the total hospital population, then it follows that it is the efficiency of the sign in the total hospital population which is of significance.

The test sign correctly identifies 86% of the suicidal patients and 81% of the non-suicidal patients (see Table I). In a population made up of all the patients (28% suicidal and 72% non-suicidal), the test sign would identify correctly 82% of the patients (or 86% of the 28% who are suicidal, plus 81% of the 72% who are non-suicidal).

To get at the efficiency of the test sign, this figure of 82% hits with the test sign should be compared with the percentage of hits obtained by calling all patients non-suicidal. The base rate figure of 28% indicates that if all of the psychiatric patients at the Menninger Hospital were to be called non-suicidal, then the description would be correct 72% of the time.

The appropriate comparison, then, shows that the test sign correctly identifies 10% more of the total hospital population than does the use of the base rate.

REFERENCE

- Appelbaum, Stephen A., and Holzman, Philip S. The color-shading response and suicide. *J. proj. Tech.*, 1962, 26, 155-161.

Answer to Dr. Jerome D. Pauker's Note on "The Color-Shading Response and Suicide"

STEPHEN A. APPELBAUM AND PHILIP S. HOLZMAN
The Menninger Foundation

The predictive efficiency of the color-shading response may be assessed in at least three ways: (1) by comparing the successful prediction of suicide with the base rate of suicide in the hospital, (2) by comparing the successful prediction of non-suicide with the base rate of non-suicide in the hospital, and (3) by comparing the successful predictions of suicide and non-suicide with the total hospital population. Doctor Pauker suggests that we should have used the third comparison for our data. The decision about which comparison to use is made on

the basis of what question is posed in the investigation, and it is analogous to the decision one makes when deciding whether or not to use a one- or two-tailed test of significance. Our purpose was to test the predictability of *suicides only* (first comparison) from the test sign. The same data could be used in the other ways as well. As it happens with these data, whichever decision one makes, one obtains an increase in predictive efficiency over the base rate through use of the color-shading determinant.

September 21, 1962

The Performance of Acute Psychotic Patients on the Index of Pathological Thinking and on Selected Signs of Idiosyncrasy on the Rorschach¹

DOUGLAS A. QUIRK
Ontario Hospital—Toronto

MARY QUARRINGTON, STEPHEN NEIGER, ALAN G. SLEMON
Toronto Psychiatric Hospital

In 1952, Watkins and Stauffacher published a weighted sign list² developed from Rapaport's (1946) listing of psychotic reactions on the Rorschach. Their purpose was to develop a psychometric device based on Rorschach observations for predicting the presence or absence of psychosis from an estimate of the amount of pathological material present in the Rorschach protocol. This scale appeared to work fairly well for the purpose for which it was developed. It was subjected to verification in studies by Powers and Hamlin (1955) and by

Pope and Jensen (1957). All studies reported adequate scoring reliability for the scale.

The present authors have been attempting for some time to develop simple, easily scored, efficient scales for the Rorschach prediction of psychosis. The scales for "common associations" (called "#(V+Vn)" and "R.I.c.", see Slemon, et al, 1961; Quirk, et al, 1961; Neiger et al, 1961c) and for "psychotic regression" (called "RSm" and "RSf", see Neiger, et al, 1961b) which were developed as a consequence of this work, appeared to be fairly efficient in distinguishing between normals and neurotics on the one hand, and chronic psychotic patients on the other.

Although neither of these latter two scales seemed to be very effective in identifying acute psychotics, particularly of the paranoid type, it still seemed possible that this group of patients could be identified on the Rorschach by means of their idiosyncratic (or "deviant") responses. It was hoped, therefore, that the Index

¹ This study was supported by funds provided under the Federal-Provincial Mental Health Grant, No. 605-5-277, of the Canadian National Health Grants programme. Grateful acknowledgement is also extended to Dr. A. B. Stokes, Head, Dept. of Psychiatry, University of Toronto for his support, and to the diverse facilities provided by the hospitals, clinics, and agencies of the Ontario Dept. of Health and by many other community and business organizations.

² Watkins & Stauffacher's (1952) "Index of Pathological Thinking" included the following items (the weighting assigned for each item is indicated in brackets following the item): Fabulized responses (.25), Fabulized combinations — spontaneously corrected or recognized (.25) — not corrected or recognized (.50), Confabulations — extreme affect loading or specificity (.50) — far-fetched elaboration (1.00) — DW (1.00), Contamination (1.00), Autistic logic (1.00), Peculiar Verbalization (.25), Queer Verbalization — usual (.50) — extreme (1.00), Vagueness (.25), Confusion (.50), Incoherence (1.00), Over-elaborate symbolism — moderate (.25) — extreme (.50), Relationship Verbalization — between two percepts (.25) — within a series of cards: (a) corrected or recognized (.25), (b) not corrected or recognized (.50), Absurd responses (1.00), Deterioration colour — colour with form (.50) — pure colour (1.00), Mangled or Distorted concepts (.25).

The signs included in the list are taken

from Rapaport's listing of reactions (1946), although brief specifications are given by Watkins & Stauffacher (1952) in the form of examples to permit greater precision of definition. There are two rules for scoring the index: (a) responses which do not qualify for any of the above scores are scored zero, and if a response qualifies for more than one of the items listed above, it is scored only once and with the appropriate item having the highest weighting, and (b) scores are not given for additional responses. When each response has received a score (0 to 1.00), the weights are summed for the entire protocol and divided by the total number of responses (R) and expressed as a percentage. The resulting percentage is designated the $\Delta\%$.

of Pathological Thinking ($\Delta\%$) would add some information in the task of discrimination, and supply a means for separating from the normal groups, those patients who were psychiatrically diagnosed as acute psychotics. Consequently, it was decided to re-evaluate Watkins and Stauffacher's $\Delta\%$, and to identify signs which could be used to facilitate this discrimination.

METHOD

Several factors were controlled in subject selection: (a) sex, (b) age, (c) language, (d) education, (e) psychiatric diagnosis, and (f) drug or physical treatment artifacts.

(a) In previous investigations the present authors had observed that sex differences tended to occur in Rorschach scores in severe psychiatric illness (see Neiger, et al, 1961b). In an attempt to control score variability somewhat, only male subjects were used in the initial phase of this investigation, and male and female groups were investigated separately during later phases.

(b) All subjects used were between the ages of 20 and 40 years.

(c) All subjects had been educated since Grade 1 in an English-speaking country.

(d) Education had been observed to have systematic effects on some Rorschach scores, and consequently, the four groups of subjects used for the initial phase of this study were selected from our files of research protocols in order to match each other for educational level. The range of educational levels for each group was from Grade 7 to University education, with the mean for each group being Grade 10. No attempt was made to control for education in the samples used for the final stage of the study.

(e) A precaution was taken to insure that none of the clinical groups contained subjects whose performance would be radically affected by drug treatments. It was specified that no patient who was receiving drug treat-

ments would be tested unless, on the day of testing, he manifested some clinical degree of the symptoms for which he was receiving the drug treatment. None of the patients had received ECT or ICT for a period of at least two weeks prior to testing, and none of them had been lobotomized.

The subjects used in the *initial* and *intermediate* phase of this study included: 20 "normal" males, 20 "neurotic" males, 20 "acute psychotic" males, and 20 "chronic psychotic" males. The four groups were matched for educational level.

The subjects used for the *final* stage of this study included: 20 male and 20 female normal subjects, 20 male and 20 female neurotic subjects, 20 male and 20 female chronic schizophrenic subjects, and 20 male and 20 female acute schizophrenic subjects. No attempt was made to control for educational level.

All psychiatric subjects were independently diagnosed by psychiatrists without disagreement and were screened clinically to exclude cases in which there was a suspicion of organic involvement.

The "normal" subjects were drawn from a large sample obtained from several large industries which made their employees available, on company time, for testing for research purposes, and from acquaintances of Rorschach students who had subjected themselves to testing for student practice. No attempt was made to screen the normal subjects for psychiatric disturbance or for indications of mental health. The "neurotic" subjects were obtained from the outpatient department of a psychiatric hospital. The psychotic patients were tested in several Ontario Hospitals.

The "acute psychotic" subjects were patients whose history of illness dated back for no more than one year from the time of testing, and who had been diagnosed as suffering from an illness in the group of schizophrenias.

The "chronic psychotic" subjects

were patients who had been in hospital for a period of time exceeding two years, who had been diagnosed as suffering from an illness in the group of schizophrenias, and who, for the most part, appeared to have rather unfavourable prognoses.

PROCEDURE

The Rorschach protocols were all individually administered according to the procedure generally followed by continental Rorschach users (Bohm, 1958; Neiger, et al, 1961a).

For the *initial* stage of this study, three sign lists are scored for each protocol. Watkins and Stauffacher's Index of Pathological Thinking (" $\Delta\%$ ") is scored by one of the investigators who appears to have adequate scoring reliability (as compared with his own performance) from a previous investigation where this scale was employed. Spot checks are made by another investigator whose delta scores, although smaller, are fairly consistent with those obtained by the first scorer. The scales for "Common Associations" [$\#(V+V_n)$] and "Psychotic Regression" (RSm) are scored by a research assistant and checked by another.

The between-groups differences on the delta scale are investigated using the standard error of mean differences, the chi-square method, and raw misplacement rates. Misplacement rates are calculated using a cut-off point of 20 and above on the Δ score³, both to observe the absolute efficiency of the delta score in prediction, and to permit the ready comparison of efficiencies of the three scales used.

³ The cut-off points for the Index of Pathological Thinking used with our data, and with Powers and Hamlin's (1955) data differ in this calculation of misplacement rates. In both cases, the cut-off points selected are the points yielding approximately the minimum number of misplacements in the data. If, instead of 10, one selects the median as the cut-off score in Powers and Hamlin's data, the misplacement rate is approximately 26%.

For the *intermediate* stage of this study, the Rorschach protocols of the same subjects as those used for the initial phase are completely scored according to an adaptation and extension of the Continental method of Rorschach analysis (Neiger, et al, 1961a). The separate signs from the Index of Pathological Thinking, as well as from our scoring are then inspected, and items having the following characteristics are identified: (a) items which appear to be most characteristic for the protocols of the "acute" psychotic group as contrasted with the other three groups; (b) items whose symptomatic values ("interpretations"), according to current Rorschach theory might be said to be associated with idiosyncrasy or unusualness of thought process, and which would usually be interpreted as reflecting bizarre or atypical symptomatic behaviour, and thus, illness; (c) items which can be defined in such a manner that they can be scored with reasonably good agreement by a great proportion of clinicians.

The items identified in this way are then listed, and the protocols each scored for the sign list. A cut-off point is selected for the sign list, and the number of subjects in each group scoring above that point determined.

For the *final* stage of this study, Rorschach tests are administered to the second group of subjects and are scored in the same manner as that described for the intermediate stage of the present investigation. The list of items for "idiosyncrasy" is scored for each subject. Misplacement rates are computed for each group, and chi-square tests are run between misplacement rates for the relevant pairs of groups. Finally, the separate items are tabulated as present or absent in order to identify the extent of the contribution of each item to the overall group identifications of the sign list.

RESULTS

The results of the *initial* stage of

TABLE I—Standard Error of Mean Differences, Chi-square, and Percentage Misplacement Rates Comparing Normals and Neurotics with Chronic and Acute Psychotics on the Index of Pathological Thinking ($\Delta\%$), Popular Responses $\#$ ($V+Vn$), a "Regression Scale" (RSm) and the Raw Number of Responses (R)

Scale	% age of Ss misplaced	Cut-off score	$\frac{Mp-Mn}{SE_{Dx}}$	p	X^2	p	N
$\Delta\%$: Present Study	33.75%	20+	3.01	.004	9.84	.01	80
Powers & Hamlin	18.0 %	10+					50
$\#$ ($V+Vn$) ^a	21.2 % ^b	6+					80
RSm	12.5 % ^b	160+					80
R	25.0 %	15+					80

^a In other work by the present authors, $\#$ ($V+Vn$) is supplied with corrections for educational effects (Quirk, et al, 1961). These corrections are *not* used here because the effects of education upon $\Delta\%$ are unknown. Instead, the groups used here are matched for educational level.

^b The misplacement rates reported here for $\#$ ($V+Vn$) and for RSm, in spite of occasional fluctuations, appear to be fairly stable in larger samples. Thus in 325 male subjects (including normals and abnormals) $\#$ ($V+Vn$) shows a misplacement rate of 22.24%, and in 305 male subjects RSm shows a misplacement rate of 13.8%.

this study, are presented in Table I. It will be seen that the delta score distinguished the groups significantly. However, it will also be seen that its misplacement rate is higher in the present data than that for either of the other two scales, and indeed than that for the raw number of Responses (R) alone.

The results of the *intermediate* stage of the study are reported in Table II. Table II presents the number of subjects scoring above the cut-off point on a (20-item) sign list developed from inspection of the Rorschach scores obtained in the acute schizophrenic group as contrasted with the other groups, and on a shorter (10-item) sign list resulting from exclusion of those items of the longer list which either have little theoretic relevance to acute schizophrenia or appeared to create difficulties for the scorer.

It will be observed that the sign lists perform with about equal efficiency across all groups.

The results of the *final* stage of this study are summarized in Tables III, IV, and V. Table III presents the raw misplacement rates⁴ for the 10-item

⁴ Since the scale has a general purpose (the identification of psychotics) as well as a special purpose (the identification of "acute"

sign list in the four male groups and the four female groups used for the validation phase of this investigation. It will be observed that there is some decline in the efficiency of the sign list in identifying acute psychotic patients, although the list continues to identify a reasonably large group of

TABLE II—Numbers of Male Subjects in the "Derivation" Group Scoring Above the Cut-off Points on a 20-item Sign List^a and a 10-item Sign List^b to Estimate Idiosyncrasy of Response in the Rorschach

Group	N	20 Item List	10 Item List
Derivation Group: Total	80	20	22
Normals	20	0	0
Neurotics	20	2	4
Chronic Psychotics	20	4	5
Acute Psychotics	20	14	13

^a Cut-off point = 4 or more.

^b Cut-off point = 2 or more.

psychotics), a "misplacement" is assumed when *any* psychotic subject scores below the cut-off point, and when a normal or a neurotic subject scores above the cut-off point. However, since the identification of chronic psychotics is achieved with adequate efficiency by means of the other scales mentioned in the text, the performance of the acute psychotics is of particular interest here. In order to keep the tables free from the complexity created by this double purpose of the scale, numbers above and below the cut-off point are reported rather than misplacements.

TABLE III—Numbers of Subjects in Male and Female Validation Groups Scoring Above the Cut-off Point (2 or more) on a 10-item Sign List to Estimate Idiosyncrasy of Response in the Rorschach

Groups	N	Number Above Cut-Off Point
Males		
Validation Group I: Total	80	20
Normals	20	0
Neurotics	20	3
Chronic Psychotics	20	9
Acute Psychotics	20	8
Females		
Validation Group II: Total	80	13
Normals	20	1
Neurotics	20	1
Chronic Psychotics	20	1
Acute Psychotics	20	10

these patients considering the difficulties encountered in identifying members of this group with the other scales referred to earlier [$\#(V+Vn)$ and RSm]⁵. It will also be observed that, although the misplacement rate appears to remain fairly low in the normal and neurotic groups, it fluctuates quite markedly in the chronic psychotic groups⁶.

Table IV presents the chi-square comparisons between relevant groups for the validation samples. It will be observed that the group discriminations, for the most part, are found to hold up significantly.

Table V presents the 10 items of the sign list with their frequencies of occurrence in the various groups. It will be observed that minus Original responses, Contaminations, and Peculiar Verbalizations carry the major

TABLE IV—Chi-square Tests of Significance of Relevant Group Comparisons for the Two Validation Samples on the 10-Item Idiosyncrasy Index

Groups Compared	χ^2
Male Validation Groups	
Normal & Neurotic vs. Chronic & Acute Psychotic	15.13*
Normal vs. Chronic Psychotic	11.60*
Normal vs. Acute Psychotic	19.55*
Chronic Psychotic vs. Acute Psychotic	.04
Female Validation Groups	
Normal & Neurotic vs. Chronic & Acute Psychotic	7.15*
Normal vs. Chronic Psychotic	0.00
Normal vs. Acute Psychotic	10.25*
Chronic Psychotic vs. Acute Psychotic	10.25*

* Significant at .01 level.

load in the sign list, while other items occur extremely infrequently.

DISCUSSION

Table I indicates that the Index of Pathological Thinking did separate the normal and neurotic groups from the "acute" and "chronic" psychotic groups to a significant degree. But it also indicates that this separation was not extremely efficient, and that there would be considerable error arising from the attempt to use this index for purposes of clinical prediction. Furthermore, the index is no more efficient, and indeed appears to be less efficient, than the other two simple indices reported [$\#(V+Vn)$ and RSm]. Also the index appears to be no more (or rather is less) efficient than the raw number of R, which can be "scored" in a fraction of the time needed for the scoring of the " $\Delta\%$ ", and which is much less subject to inter-scorer disagreement.

This last observation is particularly interesting in view of the fact that R is used in the calculation of the $\Delta\%$ in such a way that it should contribute much of its discriminating power to the discrimination performed by the delta index. That is, R forms the denominator for the delta index, so that a low R (found in "abnormals") should serve to increase

⁵ Actually, the misplacement rates are now about equal for the 10-item sign list, $\#(V+Vn)$, and RSm in the acute psychotic groups.

⁶ It appears that the female chronic schizophrenic group is responsible for this fluctuation. There seems to be a parallel between the low idiosyncrasy reaction of the female chronic schizophrenics and the essential lack of sensitivity of other measures (low form level, low R, low number of populars and others) with female chronic schizophrenics as compared with male chronic schizophrenics. (See Neiger, et al, 1961b; Neiger, et al, 1961c).

TABLE V—The Contribution of Each Item of the 10-Item Idiosyncrasy Index to the Identification of Members of the Various Groups Investigated.
Frequency of Occurrence of Each Sign in Each Group Investigated

Signs From List	Normals			Neurotics			Chron. Psy.			Acute Psy.		
	Dm ^a	Vm ^a	Vf ^a	Dm ^a	Vm ^a	Vf ^a	Dm ^a	Vm ^a	Vf ^a	Dm ^a	Vm ^a	Vf ^a
DGnt ^b	0	0	1	4	0	0	3	2	2	7	5	5
Original Minus (0—)	0	3	2	5	6	2	7	14	9	7	13	13
Unusual Sex Response	1	2	1	4	3	2	3	0	1	9	4	2
Spear, Sword, etc. R	0	0	0	0	2	0	1	2	0	6	0	0
Peculiar Content	0	0	0	0	1	0	1	2	0	2	1	0
Peculiar Verbaliz.	0	0	0	1	1	0	2	5	1	4	5	6
Neologism	0	0	0	0	0	0	1	2	1	2	1	0
Incoherence	0	0	0	0	0	0	4	0	0	5	5	0
Transformation	0	0	0	1	0	0	0	0	0	0	1	1
Inverted Card IX												
Succession	0	1	1	3	0	0	0	1	0	4	0	1

^a The column designations should read: Dm — male Derivation group, Vm — male Validation group, and Vf — female Validation group.

^b Definitions of these signs appear in the Appendix.

the delta per cent (toward the abnormal range), and a high R (tending to be found in normals) should serve to reduce the delta per cent (toward the normal range). However, no attempt was made here to test the effectiveness of the raw delta scores since there are several other problems associated with the use of the delta index which, taken together with its relatively low efficiency argue against its usefulness. The first of these problems has already been commented on at two places in the foregoing, and space does not permit a fuller account of the problems arising from the scoring of the signs in the Index of Pathological Thinking. It need only be remarked that the scale is long, complex, and is composed of many vaguely defined items. Consequently, it would not appear to satisfy the requirements of a clinical tool which is to be used widely by clinical psychologists at various levels of skill in their use of the Rorschach.

But the Index of Pathological Thinking would still be useful for some purposes if it could be shown that it performed a task which could not be performed in a simpler way. Consequently, the individual misplacements of the three scales investigated were explored to ascertain whether the delta index could be used

to contribute further identifications of subjects not correctly identified as to their group membership by means of the other scales. It was found that no subjects were correctly identified by means of the Index of Pathological Thinking who were not correctly identified by one or the other of the other two scales [$\#(V+Vn)$ or RSm], although 10 subjects were correctly identified by the delta index who were incorrectly identified by one or the other of the other two scales. At the same time, it was noted that 7 subjects were misidentified by the delta index who were correctly identified by one or the other of the two scales, and 16 subjects were misidentified by the delta index who were correctly identified by both of the other scales. These observations do not support any advantage of the delta score over the other two indices.

In the light of our introductory remarks, however, it might be hypothesized that, although the delta index demonstrates no advantage over the other simpler indices across the total group, it might show a relatively low misplacement rate in the group of "acute" psychotic patients. This hypothesis is not supported in the present data. The misplacement rates of the three indices in the group of "acute" psychotics are as follows:
 $\#(V+Vn)$..55% misplaced ($N=20$)

RSm¹.....30% misplaced

△%55% misplaced

It would appear, then, that the Index of Pathological Thinking is not particularly suitable for the purpose for which we sought to use it.

Nevertheless, the literature and clinical experience would seem to indicate that the Rorschach test can be used as a vehicle for information concerning the presence or absence of idiosyncratic or deviant ideation such as that found in patients suffering from acute psychosis. Therefore, it was decided to explore the records from the present subjects in greater detail in order to identify any signs which could be used to help in the identification of the acute records in contrast to the records from the other groups used.

The critical identification for the two sign lists resulting from this analysis was to be the group of acute psychotics. This task was found to be performed almost as well by the short list as by the long one, and it would thus seem justified to exclude the longer list from further consideration on the grounds that it requires a considerable length of time and effort to score it. Other reasons, such as the difficulty encountered in defining some of the items from the longer sign list, and the lack of theoretical relevance to the concept of idiosyncrasy of some of the items, also may be used to support the decision to exclude the longer list from further consideration at present.

The purpose of this investigation was to find a means for identifying subjects, especially acute psychotics, whose group membership was misidentified by the other two indices which had previously been developed by the present authors. There were no normal subjects misidentified by both # (V+Vn) and RSm. None of the three normal subjects misidentified by one or the other of # (V+Vn) or RSm were misidentified by the 10-item list. The one neurotic subject misidentified by both # (V+

Vn) and RSm was correctly identified by the 10-item list, and the three other neurotics misidentified by one or the other of # (V+Vn) or RSm were correctly identified by the 10-item list. Of the chronic psychotic patients, two had been misidentified by one of the scales [# (V+Vn)], and one of these was correctly identified by the 10-item list⁷. Finally, of the acute psychotic group, four of the paranoid patients had been misidentified by both # (V+Vn) and RSm, and three of these subjects were correctly identified by the 10-item list; nine of the patients had been misidentified by one or the other of # (V+Vn) or RSm, and one of these was correctly identified by the 10-item list; and four of the patients identified incorrectly by the 10-item list were correctly identified by both of the other two scales [# (V+Vn) and RSm]. It will be seen that in the first (derivation) group studied the 10-item list seemed potentially useful for increasing the predictive efficiency of the previously developed two scales.

In addition, the validity of the 10-item scale, considered by itself, appeared to be supported by its ability to discriminate significantly among all the groups used in the initial stage of the study, with the exception of the normals versus the neurotics (and these groups, of course, were not predicted to differ). The chi-square contrasts for the group comparisons are:

Normal and neurotic versus chronic and acute psychotic:

$$X^2=12.28, p < .001$$

Normal versus chronic psychotic:

$$X^2= 5.71, p < .02$$

Normal versus acute psychotic:

$$X^2=19.26, p < .001$$

Chronic versus acute psychotic:

$$X^2= 6.47, p < .02$$

Nevertheless, the utility of this

⁷ The concept of "misplacement" with respect to chronic schizophrenics refers to the fact that this is a scale aimed at the identification of psychotic subjects. A "correctly" identified chronic schizophrenic scores above the cut-off point.

scale could not be assessed alone on the basis of the groups from which it was derived. Further investigation of the sign list was required to assess the consistency with which the list could facilitate group discrimination. It also remained to be seen if the scale which was derived from male groups of subjects, could perform with satisfactory efficiency in female groups.

The findings with the validation sample support the conclusion that relatively short-term psychotic patients are identifiable (at least) about half the time on the basis of their Rorschach responses, and that the face-value behavioural character of the responses used here for this purpose would suggest that this can be done on the basis of these patients' idiosyncratic behaviour. Furthermore, examination of the validation groups' performances on the 10-item index and on $\#(V+V_n)$ and the Regression Scale (RSm or RSf) reveals the following misidentifications of subjects in the two "acute psychotic" groups: Among the male acute psychotics, 3 of the 9 patients misidentified by both $\#(V+V_n)$ and RSm were correctly identified by the 10-item index and 3 of the 7 patients misidentified by one or the other of $\#(V+V_n)$ or RSm but not by both, were correctly identified by the 10-item index. Similarly, among the female "acute psychotics", 9 of the 14 patients misidentified by both $\#(V+V_n)$ and RSf were correctly identified by the 10-item list, and neither of the 2 patients misidentified by one or the other of $\#(V+V_n)$ or RSf was correctly identified by the 10-item list. It will be noted that the usefulness of this sign list, in facilitating identification of acute patients misidentified by $\#(V+V_n)$ and RS, which was observed in the derivation sample holds up in the validation sample. It is to be hoped that eventually dimensional prediction will be possible. Plotting a subject's position on several indices whose relationships to one another in the task of prediction are known,

might permit an overall increase in the efficiency of prediction by providing differential information relevant to differential diagnosis. This task, however, remains for the future.

In the second place, it may be stated that, in spite of the apparent fluctuations in the usefulness of some idiosyncratic signs due to their low frequency of occurrence in any groups, a relatively few Rorschach signs are needed in order to attain a relatively satisfactory consistency in the efficiency with which acute psychotic subjects can be identified by means of their Rorschachs. Again, it must also be stated that there is rather slight consistency in the efficiency of identifying nonacute psychotics by means of these same signs.

In the third place, it should be noted that some of the signs used in the present investigation could probably be excluded from the present list without impairing its efficiency noticeably, due to the extremely low frequency of occurrence of these signs in all groups, or to their relative lack of power in discriminating between the groups. Thus, for example, Table V would seem to suggest that the sign list might be reduced to a list of 3 to 6 items in length, namely, DGnt, O-, Peculiar Verbalization, Neologism, Incoherence, and/or Unusual Sexual Content.

It is clear, however, that the sign list as presently constituted, is not suitable for clinical use without the support of other indices. Further work remains to be done by way of replication and refinement before the validity and usefulness of such a sign list may be said to have been demonstrated.

SUMMARY

The problem of this investigation was to find a method by means of which acute psychotic patients could be distinguished from normal and neurotic subjects. Watkins and Stauffer's Index of Pathological Thinking in the Rorschach was first investi-

gated to determine whether it would fulfill this function. Although the delta score was found to distinguish psychotic and nonpsychotic groups significantly, it did not add an appreciable amount of information to the prediction of individual subjects' group membership over that supplied by two other simple and efficient indices which had been developed previously to aid clinical prediction. Consequently, an attempt was made to identify other Rorschach signs occurring with relatively higher frequencies in the records of acute psychotic subjects than in the records of other subjects. Twenty such signs were found, and ten of these were listed as having theoretic relevance to the concept of idiosyncrasy. These 10

signs were then tested against the 20 normals, 20 neurotics, 20 chronic psychotics and 20 acute psychotics (all males) from an examination of whose protocols the signs were derived.

The 10-item sign list was found to work reasonably efficiently on its own derivation group, and was then applied to two new groups (one male, one female) composed in the same manner and with the same numbers as the (all male) derivation group. In the validation groups, the group comparisons between the critical clinical subgroups remained statistically significant, and about half of the acute psychotics were correctly identified as to their group membership. Nevertheless, the sign list requires further replication and refinement.

APPENDIX

Below are the definitions for 10 signs and the names of 10 other signs identified by inspection as occurring more frequently in acute psychotics' protocols than in the protocols of other subjects. The first 10 signs, which are listed with their definitions, form the 10-item sign list referred to in this paper, and they were selected from the 20-signs listed below as having some relevance to the concept of idiosyncrasy as it is viewed through Rorschach theory.

1. *DGnt: Contaminated secondary whole responses* (Rorschach, 1942). These are whole responses created by the condensation of details which, by themselves, are usually acceptable (good form level), but either the way of organizing them is *perceptually* absurd, or else the parts are logically incompatible in that concept. Four types of DGnt are recognized: (a) one area is used for more than one concept in the same interpretation (e.g., "A butterfly in the shape of all organs"); (b) drastic incompatibility in size without the realization of the incompatibility (e.g., VIII, "Two cats climbing a huge mountain"); (c) drastic incompatibility in the location of the parts used (e.g., III, all black

and centre red, "Two diplomats with their hearts joined"); (d) incompatibility of the detail concepts in the total concept (e.g., VIII, "A fur coat (bottom D) out of which grows one tree and two animals. They are all grown together" (whole)).

2. *O—: Original minus responses* (Rorschach, 1942). Original minus is a response in which the concept is highly unusual for that location and it does not "fit" the blot area used for it. In order to reduce the number of problems in communicating this category, the Hertz (1951) tables were used as the standard reference, both for originality (responses occurring no more frequently than once in a 100 protocols) and for form level. When a response occurred which was not listed in Hertz's tables, one of the investigators made the decision.

3. *Unusual Sexual Response:* (Rapaport, 1946). A sexual response may be unusual either (a) because of its location, or (b) because of its content, or (c) due to the manner of its verbalization.

(a) In our opinion there are only six areas on the Rorschach cards that are acceptable for external genital responses. These areas are:

1. Centre bottom red on II (female)
2. Top centre, black on II (male)
3. Small top centre part in shading on IV (female)
4. Upper one-third of VI (male)
5. Bottom centre indentation on VI (female)
6. Centre bottom on VII (female)

If sex responses are given to any *other area*, they are considered unusual.

(b) and (c) Similarly, references to the *sex act itself*, *peculiar references* to anatomical aspects of the sexual process ("germ joining ovum"), and sexual content marked by *blunt vulgarity* are considered unusual. Finally, hesitation, vagueness, and embarrassment accompanying even usual sexual responses are given credit as Unusual Sexual Responses.

4. *Spear, Sword, etc., Content:* (Merei, 1953). The subject uses as a content for any of his responses the concepts: spear, sword, lance, knife, arrow, cutlass, sabre.

5. *Peculiar Content:* (Rapaport, 1946). A response of two or more parts which it would seem extremely peculiar or strange to bring together in just that way and form (e.g. "Various genitals on display").

6. *Peculiar Verbalization:* (Rapaport, 1946). Use of a phrase (two or more words) in a peculiar context, where the aim is *not merely* to appear clever (i.e., a "big word" which is wrongly used), but where the result is to drag in an unnecessary and only partly related idea (e.g., "Medically speaking, this would be a picture of lungs").

7. *Neologism:* (Bleuler, 1920). Subject formulates a new word to express a concept. This is usually the result of a conceptual condensation (e.g., "poke stick" for billiard cue, or "fantabulous").

8. *Incoherence:* (Rapaport, 1946). Subject strings words and phrases together, usually with the first phrases as an association to a response, with no order or necessary connection of

the remainder of the association to the response, although the associations may bear some connection to one another in the series (e.g., "A lovely garden . . . blue . . . like my sister's coat . . . she's a lovely woman . . . and her husband . . . flows into a nice amber pink").

9. *Transformation:* (Holt, 1956; Neiger, 1956). Subject verbalizes explicitly that the blot or the thing seen in the blot is changing (transforming itself) right before his eyes (e.g., "A bat. Wait, it's changing. It's becoming something different. It's changing into a wasp").

10. *Card IX, Inverted Succession:* (Neiger, et al, 1961a). The subject inverts the succession of location scores on Card IX. The usual succession of location scores is G (whole) response(s), then D (large detail) response(s), then perhaps Dd (small detail) response(s), with the succession sometimes repeating itself after the card has been turned to a new position. In inverted succession, the subject (a) progresses from D response(s) to G response(s), or from Dd response(s) to D response(s); or (b) produces no response(s) other than Dd.

11. *Contamination:* (Rapaport, 1946; Watkins & Stauffacher, 1952).

12. *Sex Concern:* (Neiger, et al, 1961a).

13. *Colour Claim:* (Merei, in Neiger, 1956).

14. *Closure:* (Quirk, in Neiger, et al, 1961a).

15. *Confabulation Reaction:* (Bohm, 1958).

16. *Movement in a Double Sense:* (Zulliger, 1941).

17. *Abstraction and Symbolism:* (Rapaport, 1946).

18. *Self Criticism:* (Frankel & Benjamin, 1934).

19. *Confusion:* (Rapaport, 1946; Watkins & Stauffacher, 1952).

20. *Vagueness:* (Rapaport, 1946; Watkins & Stauffacher, 1952).

REFERENCES

- Bleuler, E. *Textbook in psychiatry*. (2nd ed.) Berlin: Springer, 1920.
- Bohm, E. *A textbook in Rorschach test diagnosis*. New York: Grune & Stratton, 1958.
- Frankel, F., & Benjamin, D. Die Kritik der Versuchsperson beim Rorschach'schen Formdeutversuch. *Schweizer Archiv. f. Neur. u. Psychiatrie*, 1934, 33, 9-14.
- Hertz, Marguerite R. *Frequency tables for scoring responses to the Rorschach inkblot test*. Western Reserve University, 1951.
- Holt, R. R. Primary and secondary processes in Rorschach responses. *J. proj. Tech.*, 1956, 20, 14-25.
- Merei, F. *Der Aufforderungscharakter der Rorschach-Tafeln*. Translation and comments by Neiger, S., Innsbruck, Arbeitsgemeinschaft f. Rorschach u. a. proj. Testmethoden, 1953.
- Neiger, S. Specific reactions. Part II of *Introduction to the Rorschach psychodiagnostic*. Mimeo. Toronto Psychiatric Hospital, Ontario Department of Health, 1956.
- Neiger, S., Quirk, D.A., & Slemon, A. G. *Manual of the Rorschach psychodiagnostic*. Toronto, author, 1961.
- Neiger, S., Slemon, A. G., & Quirk, D. A. Scales of psychotic regression from the Rorschach Test. Toronto, author, 1961.
- Neiger, S., Slemon, A. G. & Quirk, D. A. Rorschach indices of common associations from popular responses. Part III. Validation and cross validation of indices. Toronto, author, 1961.
- Pope, B., & Jensen, A. R. The Rorschach as an index of pathological thinking. *J. Proj. Tech.*, 1957, 21, 54-62.
- Powers, W. T., & Hamlin, R. M. Relationship between diagnostic category and deviant verbalizations on the Rorschach. *J. consult. Psychol.*, 1955, 19, 120-124.
- Quirk, D. A., Slemon, A. G., & Neiger, S. Rorschach indices of common associations from popular responses. Part II. Preliminary standardization of indices. Toronto, author, 1961.
- Rapaport, D. *Diagnostic psychological testing*. Vol. II Chicago Year Book Publishers, 1946.
- Rorschach, H. *Psychodiagnostics*. Berne: Huber, 1942.
- Slemon, A. G., Neiger, S., & Quirk, D. A. Rorschach indices of common associations from popular responses. Part I. Development and selection of indices. Toronto, author, 1961.
- Watkins, J. G., & Stauffacher, J. C. An index of pathological thinking in the Rorschach. *J. Proj. Tech.*, 1952, 16, 276-286.
- Zullinger, H. *Behn-Rorschach-Test*. Berne: Huber, 1941.

Received December 6, 1961

Revision received June 18, 1962

Some Behavioral Correlates of the Rorschach Experience-Balance¹

MELVIN ROSENTHAL
Boston University School of Medicine

The translation of Rorschach performance into psychological language and concepts is an ongoing process that is aided by validation data obtained through extra test experimentation. The following study presents data bearing upon certain specific assumptions concerning the interpretative significance of the experience-balance.

The concept of the experience-balance expresses the relationship between the quantity of human movement responses and the sum of the weighted color responses, the ratio $M:\Sigma C$. The importance of the experience-balance has long been recognized as representing one of the basic personality dimensions (Beck, 1945; Klopfer, 1946; Rorschach, 1942).

There are conflicting viewpoints related to the problem of experimental validation of the Rorschach test. Previous attempts at research with the experience-balance consist of; 1) typological studies, which attempted to relate Rorschach's experience types to other typological systems (Guilford, 1934; Hertz, 1935); 2) clinical studies, which comprise the bulk of past validation work, utilizing clinical entities and diagnostic groups as criteria (Phillips, 1953); and 3) experimental studies which compare Rorschach variables with independent and experimentally controlled behavioral criteria (Singer *et al*, 1950; Williams, 1952; Palmer, 1962). The general plan of this experiment was to compare variation of performance in an independent, standardized, situation as a function of variation in the

Rorschach factor, the experience-balance.

A review of the Rorschach literature revealed some common consensual interpretations. From these a broadly stated assumption was constructed.

Human movement responses reflect a tendency to emphasize behaviorally mental activity whereas color responses reflect a tendency to emphasize behaviorally motor activity.

Two behavioral predictions were derived from this assumption:

- 1) People who give a preponderance of the human movement determined responses to the Rorschach test delay motor activity longer in approaching a novel situation than do people who give a preponderance of color determined responses.
- 2) People who give a preponderance of human movement determined responses to the Rorschach test are motorically less active in a novel situation than are people who give a preponderance of color determined responses.

METHOD

Subjects: Variation in the experience-balance served as the primary criterion for the selection of the subjects. Median cutting points were determined for both M and C using eighty-six Rorschach protocols. The obtained median values were: $M=3$, $C=2$. The M group consisted of those subjects in whom M was above three and C at or below zero. The C group consisted of those subjects in whom C was above two and M at or below three. Wherever possible a subject was excluded from either group if one score, either M or C, occurred in the cell of the median value (Table 1.)

The eighty-six Rorschach protocols yielded two groups of ten subjects each. Both groups with one exception consisted of white, American-born students drawn from psychology

¹ This paper is a condensation of a part of a doctoral dissertation done at Boston University. The author wishes to express his appreciation to Drs. A. William Hire, Austin W. Berkeley, and Leo J. Reyna.

TABLE I—Age, Sex, Wonderlic Score, R, and Experience Balance For The M and C Groups

Subject	Age	Sex	Wonderlic Score	Total R	M:C
<i>M Group</i>					
1	19	F	26	15	1:2
2	22	M	29	28	7:1
3	21	F	28	16	4:2
4	19	F	25	33	6:2
5	22	M	25	13	6:1.5
6	21	M	33	11	4:5
7	25	M	32	19	9:2
8	20	F	32	10	7:0
9	23	M	30	26	10:1.5
10	40	M	26	49	4:5
<i>C Group</i>					
1	19	M	28	34	3:6.5
2	38	M	30	16	2:6
3	37	F	26	24	0:4.5
4	25	M	23	46	2:3.5
5	32	M	33	37	2:4.5
6	21	M	24	19	2:1.5
7	21	F	28	19	1:5.5
8	35	M	28	28	1:5
9	22	F	28	25	2:3.5
10	21	F	32	26	1:4

classes conducted at a local university. One subject was obtained from a group of student nurses in training at a local hospital. No subject with a history of, or treatment for psychiatric or neurological difficulties was used. Any subject with a record of previous Rorschach examination was eliminated from the experiment. The hypothesis that the two samples were drawn from a common population with respect to age, sex, intelligence, and total number of Rorschach responses was found tenable (Table I).

Procedure: The subjects were seen in two sessions. The first session consisted of the administration of the group Rorschach test. The second session was made up of the individual administration of the Stick tasks and the Wonderlic Personnel Inventory, Form A, which was used to derive an estimate of intellectual level (Wonderlic, 1945).

Stick Test: A modified version of the Katona (1940) Match Stick Tests was employed as the independent, standardized behavior situation. This test was selected for the following reasons: (a) it offered the subjects an

opportunity to behave in a manner compatible with both aspects of the assumption and its implications for behavior; (b) it was possible to vary and apply the material extensively; and (c) these tasks proved difficult enough in a pilot study to allow for extended behavioral observations. Nine tasks were used. They were arranged in an ascending order of difficulty as determined by the pilot study. Reaction time, defined as the time elapsing between the last word of the instructions to the time the first move was made, served as the measure of delay. A move was defined as any change of position of a stick either from the arrangement placed before the subject, or from an arrangement he had devised in attempting solution. The number of moves served as the measure of motor activity.

The behavioral implications, deduced from the assumption, were translated into operational hypotheses.

1. The reaction time for M group will be significantly longer than the reaction time for C group.
2. The M group will make signifi-

cantly fewer moves than will the C group.

RESULTS

Statistics: This study was designed so that both operational hypotheses could be tested by analysis of variance for repeated measurements of the same subjects (Edwards, 1951). In order to make tenable the assumption of normality and homogeneity of variance the reaction times were transformed to natural logarithms. The number of moves made on any given task were dealt with by the ratio, Rate of Moves, given by the formula, moves/time units (15 seconds) was used, since all subjects did not use an equal amount of time in working on the tasks. Solving the problem reduced the total time, and made it necessary to utilize a measure which took time into consideration. A subject's score for a single trial was determined by dividing the total number of moves on that trial by the number of fifteen second units used in attempting solution. Thus, eight moves in five time units resulted in the ratio $8/5 = 1.600$.

The null hypothesis was invoked for each variable and tested against the classes of alternatives which state that: 1) the mean log reaction time for the M group is larger than the mean log reaction time for the C group; and 2) the mean rate of moves for the M group is lower than the mean rate of moves for the C group.

From the results of the analyses of variance, cited in Tables II and III, acceptance of both stated alternatives is indicated. The between groups F test for log reaction time yields the value 11.616, which is significant at $<.01$ level. The value of F for rate of moves was 29.982, significant at $<.001$ level. Tests of homogeneity of variance were performed to ascertain whether the highly significant values of F could be attributed to mean differences, and the null hypothesis was accepted in both instances.

Statistical analysis of the data on number of solutions was made. The null hypothesis was tested for each of the nine trials by means of the Chi-square test. Yates' correction for continuity in all nine instances was utilized. The results of the Chi-square analyses are presented in Table IV, and indicate acceptance of the null hypothesis in all nine instances.

It would seem that while the M and C groups clearly manifest differential modes of approach, they do not differ with respect to efficiency of solution.

DISCUSSION

The evidence indicates support for the hypothesis that subjects who respond to the Rorschach test with human movement percepts predominating over color percepts are more delayed in their behavior and are physically less active than color-predominant perceivers. In light of the supplemental finding that the groups did

TABLE II—Analysis of Variance of Log Reaction Time for Two Groups of Subjects Tested on Nine Trials

Source of Variation	df	Mean Square	F	P
Between groups	1	10.420	11.616	$<.01$
Between subjects in the same group	18	.897		
Total	19			

TABLE III—Analysis of Variance of Rate of Moves for Two Groups of Subjects Tested on Nine Trials

Source of Variation	df	Mean Square	F	P
Between groups	1	162.890	29.982	$<.001$
Between subjects in the same group	18	5.433		
Total	19			

TABLE IV—Chi-square Tests for Number of Solutions Between the M and C Groups

Task	Number of Solutions by Group		χ^2	p
	M	C		
I	8	4	1.875	.10
II	9	9	.556	.30
III	8	7	0	.99
IV	2	2	.313	.50
V	6	5	0	.99
VI	5	4	0	.99
VII	4	1	1.067	.30
VIII	3	4	.879	.30
IX	5	0	0	(9)

not differ in efficiency, it seems appropriate to interpret the results as indicating that the experience-balance reveals characteristic response tendencies inherent in the "mode of approach" concept.

Movement-preponderant people may be characterized as being cautious and deliberate in their approach to novel situations. On the other hand color-preponderant people appear to jump into things with relatively little delay or caution. Once actively involved in their attempts to resolve the situation, they tend to be far more active motorically than movement-preponderant individuals.

It would seem premature to conclude that movement-predominance indicates a different personality type in contrast to color-predominance. Individuals may switch response tendencies exhibiting flexibility in mode of approach consistent with task requirements. Furthermore, the availability of alternative response tendencies and the ability to shift may be different in patients than in non-patients. Perhaps it is that as individuals they have developed a preference for a particular mode of approach. Those responses consistent with the set are the responses upon which they have come to rely and express most easily. When confronted with a free choice situation, these response tendencies are manifested to the relative exclusion of others.

The behavior of one subject in the M group was particularly revealing in this respect. After working deliber-

ately on the first three tasks, which he solved with apparent ease, he seemed to encounter considerable difficulty on the following task. He made no moves for the first two minutes at which point he asked, "Is it all right to experiment?" When told he could proceed in any way he desired, he made nineteen moves, in a random, trial and error fashion, during the next minute. During this time, his behavior resembled closely the behavior manifested by C group subjects. In the last two minutes of the allotted time, and on succeeding tasks, he reverted to his more characteristic behavior, apparently disappointed that the change in approach had not yielded a successful completion of the task.

To conclude, some individuals prefer one mode of approach to novel experiences rather than another. In this sense individuals who give a preponderance of human movement responses to the Rorschach test may be said to prefer a mode of approach characterized by deliberation, caution and little motor activity. Individuals who give a preponderance of color determined responses may be said to prefer a mode of approach characterized by little delay and a high degree of motor activity.

Thus, the findings of this study confirm the validity of some aspects of the Rorschach experience-balance. In addition, the production of comparable achievement scores emphasizes the importance of process differences which are frequently overlooked when

performance style is not taken into account.

REFERENCES

- Beck, S. J. *Rorschach's test. Volume II: A variety of personality pictures*. New York: Grune & Stratton, 1945.
- Edwards, A. L. *Experimental design in psychological research*. New York: Rinehart, 1951.
- Guilford, J. P. "Introversion - extroversion." *Psychol. Bull.*, 1934, *31*, 331-354.
- Hertz, Marguerite R. "The Rorschach ink-blot test: Historical summary." *Psychol. Bull.*, 1935, *32*, 33-66.
- Katona, G. *Organizing and memorizing*. Morningside Heights, New York: Columbia Univer., 1940.
- Klopfer, B. and Kelley, D. M. *The Rorschach technique*. Yonkers-on-Hudson, New York: World Book, 1946.
- McNemar, Q. *Psychological statistics*. New York: John Wiley, 1949.
- Palmer, J. O., Lustgarten, B. J. The Prediction of TAT structure as a test of Rorschach's Experience-Balance. *J. proj. Tech.*, 1962, *26*, 212-220.
- Phillips, L. and Smith, J. C. *Rorschach interpretation: Advanced technique*. New York: Grune & Stratton, 1953.
- Rorschach, H. *Psychodiagnostics: A. Diagnostic test based on perception*. (Translated by Lemkau, P., and Kronenberg, B.) New York: Grune & Stratton, 1942.
- Singer, J. L., Meltzoff, J., & Goldman, G. D. "Rorschach movement responses following inhibition and hyperactivity." *J. consult. Psychol.*, 1950, *14*, 283-286.
- Williams, M. "The Experimental validation of the Rorschach test. I. Experimental correlations." *Amer. J. Orthopsychiat.*, 1952, *22*, 749-754.
- Wonderlic, E. F. *Wonderlic Personnel Test manual*. Glencoe, Illinois; E. F. Wonderlic, 1945.

Received April 3, 1962

The Performance of Mentally Defective Children on the Color Pyramid Test¹

K. WARNER SCHAIE
University of Nebraska

Personality assessment procedures are typically restricted in their usefulness with a mentally defective population due to the fact that these procedures capitalize heavily on verbal behavior. It is the rare exception (and then generally only for Ss at the borderline level of intelligence) that a self-report technique is responded to in other than a random fashion. When projective techniques such as the Rorschach or TAT are used, meaningful responses may be obtained, but records are generally characterized by poverty of content and by a stereotypy which seems to defy interpretations going beyond broad generalizations which could be provided with greater parsimony by direct behavior observations. Non-verbal methods such as doll-play techniques may be more applicable but raise the question whether the relative complexity of verbal instructions and examiner-client relationship can lead to meaningful results with defective Ss.

It is therefore of interest to consider a technique which is of minimum complexity both in terms of the task involved and in the instructions required. Such material needs to be non-verbal, but in other populations must have been clearly related to personality attributes and behavior variables. These conditions are met by several of the techniques involving color preference or color structuring.² While a

good deal is known about the color preferences and color connotations of normal subjects (Guilford, 1959; Norman and Scott, 1952), little work has been done with mentally defective subjects in this area. Search of the literature yielded only one study (Stacey and Reynolds, 1953) which described the preferences of sub-normal children, but this study does not provide normal controls nor does it employ a method which has been related to other behavior variables.

The present study is concerned with the analysis of the response of a group of mentally defective children to the Color Pyramid Test (CPT), a formalized color arrangement task which seems to fit the above criteria. Since several personality and behavior correlates of this test have been investigated (Heiss and Hiltmann, 1951; Schaie, 1962, 1963), it may be possible to generate hypotheses and predictions about such attributes in mentally defective children on the basis of their CPT performance.

The Color Pyramid Test is described in greater detail elsewhere (Heiss and Hiltmann, 1951; Schaie, 1962b). It will suffice here to indicate that in this test the subject is presented with a selection of 1" squares in 24 different hues (including all the major colors of the spectrum plus black, gray, brown, and white) and is asked to fill the fifteen fields of a pyramid, a reproduction of which is shown in reduced size in Figure 1. The subject is instructed to make the pyramid as "pretty" as possible using any colors he wishes and arranging them at his pleasure. Three repetitions are required; then the instruction is changed to require construction of an "ugly" pyramid, again with three

¹The author is grateful to the Research Council of the University of Nebraska for financial assistance of this research, to John D. Burchard who participated in the data collection, to Stanford Gilbert for assistance with the data analysis, and to M. E. Wyant, Superintendent of Beatrice State Home, for permission to test subjects in his institution.

²The rationale for the use of response to color as a tool for personality assessment is considered elsewhere (Schaie, 1961a; 1961b).

trials. The frequency with which each hue is used is tabulated and the pattern of colors and frequency of use of each color is hypothesized to yield information relevant to the individual's control of affect as well as other personality attributes.

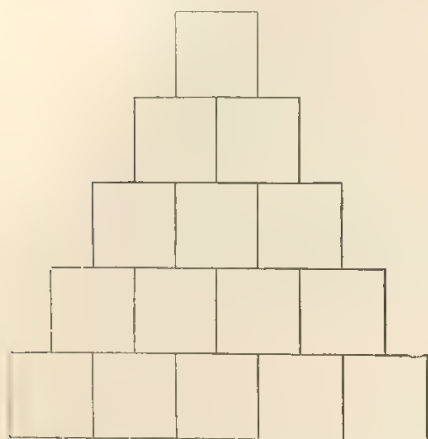


FIGURE 1—Reduced Reproduction of the Pyramid Form.

Since evidence is accumulating that CPT responses are subject to developmental shifts (Heiss and Hiltmann, 1951; Karl, 1956; Schaie, 1961c) it is appropriate to examine the hypothesis that the test will be sensitive to and will detect developmental retardation. Furthermore, if the CPT is to be used for the description of affect and other personality attributes in a retarded population it is necessary to identify those variables in which retardates differ from a normal population on the basis of their retardation, and those variables which may be indicative of emotional disturbance. It is the purpose of this paper to examine these questions by comparing the performance of a group of mental defectives with a sample of normal children.

PROCEDURE

The Color Pyramid Test was administered according to standard instructions to 29 girls and 29 boys at

a State Home for the Mentally Retarded. Ss were selected by inspecting the alphabetical roster of patients and selecting all those whose record showed no apparent organic pathology or severe emotional disturbance and whose mental age (as measured by the Stanford-Binet) was 5 years or greater.³ All Ss were native-born whites. The male Ss had a mean C.A. of 13.8 years with a S.D. of 3.1. Their mean M.A. was 7.3 years with a S.D. of 1.7 and their mean IQ was 51.3 with a S.D. of 6.4. Female Ss had a mean C.A. of 13.8 (S.D. = 2.4); a mean M.A. of 7.1 years (S.D. = 1.3) and a mean IQ of 51.8 (S.D. = 5.4). Three Ss were tested at one time, in such a manner that they could not transmit response cues to one another.

Raw scores for each of the 24 hues were combined for the 10 major colors into which the hues are conventionally grouped and were then converted into Sten scores utilizing norms for normal school children (Schaie, 1961c). These tables, providing separate norms for boys and girls in the Lincoln, Nebraska, public school system, were constructed using random samples of 50 school children for each 2 year age interval from 6 to 18 years of age. The tables were entered once with the S's chronological age (the resulting Sten scores to be referred to as CA scores) and once using the S's mental age (to be referred to as MA scores).

RESULTS

The first question to be investigated was the hypothesis that there might be no differences in CPT performance between normal and retarded children. Means and S.D.s were computed for the CPT-CA scores separately for boys and girls in our retarded sample and the *t* test was utilized to examine the significance of the difference be-

³The lower age limit was imposed, as normal subjects below five years of age frequently tend to see their task as filling the pyramid without attention being paid to either colors or form, the primary stimulus variables.

tween the obtained means and those of the normal sample for each of the 10 colors⁴. Since the data for the normal group are in Sten score form, means and sigmas for the normal Ss on all variables were 5.5 and 2.0 respectively, based on a sample size of 50 subjects. Table I gives the means and standard deviations for each of the CPT-CA scoring variables, as well as the results of the *t* test comparisons with respect to the above hypothesis.

Examination of Table I shows a considerable number of significant discrepancies between the mental defectives and the norm group. On the pretty pyramids, both male and female Ss were significantly high on Purple and significantly low on Blue. The female Ss alone were also significantly high on Orange and the male Ss were high on Brown⁵. On the ugly pyramids, both males and females were significantly high on Yellow, Blue, and White, but significantly low on Black. The males alone were also significantly low on Brown and Gray.

These results, particularly those which hold for both sexes, appear quite consistent with the interpretive rationale suggested by Heiss (1952), as well as some German experimental findings with subnormal children (Heiss and Hiltmann, 1951). Following these authors, the high Brown is considered evidence of developmental disturbance, the high Orange of excessive externalization of affect, the high Purple of emotional disturbance, and

TABLE I—Means and Standard Deviations for the CPT-CA Scores of the Mentally Defective Ss and Results of the *t* Tests of Differences of the Means with Those of the Norm Group of Normal Ss

Colors	Pretty Pyramids			Ugly Pyramids			Females		
	Mean	S.D.	<i>t</i>	Mean	S.D.	<i>t</i>	Mean	S.D.	<i>t</i>
Red	5.93	1.50	1.18	6.14	1.83	1.45	5.00	1.78	1.28
Orange	6.00	1.71	1.18	6.62	1.40	2.92**	5.62	1.32	2.30
Yellow	5.07	1.64	1.03	4.97	1.18	1.48	6.18	1.55	2.28*
Green	5.59	1.66	.22	5.03	1.48	1.19	5.00	1.17	1.11
Blue	4.35	1.16	2.68**	4.83	1.39	1.75	5.00	1.78	2.07**
Purple	7.14	1.13	4.66**	7.00	1.41	3.90**	5.55	1.99	1.11
Brown	7.10	1.51	3.97**	6.07	1.13	1.62	4.86	2.01	1.39
White	5.58	2.17	.16	5.34	1.14	1.5	4.82	1.65	3.34**
Gray	5.07	1.91	1.26	5.72	1.28	.60	5.32	1.76	.28
Black	5.66	1.34	.42	5.17	1.14	.93	5.62	1.66	2.27**

*Significant at or beyond the 5% level of confidence.

**Significant at or beyond the 1% level of confidence.

⁴Since multiple comparisons are involved, the analysis of variance was first used to test the hypothesis of overall differences. The overall differences between normals and defectives, simple interaction between groups and colors, triple interaction between groups, instructions and colors, and the fourfold interaction between groups, instructions, sex and colors were all significant at or beyond the 1% level of confidence. The analysis of variance data are not reported in greater detail since primary interest is in the mean comparisons.

⁵The female Ss show a trend for a high score on Brown also. The mean difference here, however, was significant at only the 10% level of confidence.

the low Blue of failure to develop cognitive means for the control of affect. The responses to the ugly pyramid instruction are said to reflect rejection of affective models. Thus, again using the test authors' rationale, the high Yellow reflects rejection of emotional investment in object relationships, the high Blue rejection of cognitive or rational modes of affect control, the high White the rejection of internalization or withdrawal and the low Black lack of sensitivity to the inappropriateness of depressive modes of affect control.

The question arises, however, whether the above attributes are indeed dominant personality attributes of the mentally retarded population, or whether they appear as an artifact of mental age and thus indicative of developmental retardation at a given chronological age. The latter hypothesis seems a reasonable one when one observes the gradients of CPT performance for normally developing children (Schaie, 1961c). The frequency for Blue for example, shows a positive age gradient, suggesting that the lower mean for the defective group may indicate performance typical of a younger age level. Similarly, Purple shows a negative age gradient for both sexes, while the gradients for the other variables showing significant age differences are somewhat more complex.

The above considerations suggest that a formal distinction needs to be made between those CPT variables on which the defectives differ from the normals due to intellectual retardation; and the variables which reflect personality differences not directly related to such retardation. To permit such an analysis, all raw scores were re-entered in the tables of norms using the S's mental age as the reference age. The resulting CPT-MA score means were then compared with the CPT-CA score means.

If differences in CPT performance by the defective group are indeed due to intellectual retardation, then the

CPT-MA scores should be closer to the population mean than the CPT-CA scores. A one-tailed t test is therefore appropriate, with the null-hypothesis depending upon whether the CPT-CA mean is above or below the population mean. Thus, if $m_{CA} > 5.5$, then $H_0: m_{CA} \geq m_{MA}$ with $H_1: m_{CA} < m_{MA}$; and if $m_{CA} < 5.5$, then $H_0: m_{CA} \leq m_{MA}$ with $H_1: m_{CA} > m_{MA}$. Results of the t tests are given in Table II and it appears that more than half of the variables are affected by the re-scoring procedure. On the Pretty Pyramids, means for both male and female Ss shift significantly in the predicted direction for Orange, Blue, Purple, and Gray. In addition significant shifts occur for male Ss alone on Blue, and for female Ss alone on Red, Green, and Black. On the Ugly Pyramids, significant shifts in the predicted direction occur for both male and female Ss on Red, Green, White, and Gray; for male Ss alone on Blue, Purple, and Brown; and for female Ss alone on Yellow. It may thus be concluded that performance on these CPT variables is at least in part affected by the Ss' level of intellectual development.

The results of the mental age correction will, of course, seriously affect possible interpretations of personality attribute differences made on the basis of the original scores. For those variables where significant shifts occurred, it is therefore necessary to examine the CPT-MA score means to see whether the correction has eliminated the previously noted significant differences from the normal comparison population, or whether in spite of such significant shift in the direction of the population mean, a significant difference remains. It was therefore necessary to repeat the analysis summarized in Table I, but using the CPT-MA instead of the CPT-CA scores for the comparisons with the norm group. Means, Standard Deviations, and results of the appropriate t tests for this further analysis are listed in Table III.

It now appears that originally reported differences for the male Ss on

TABLE II—Mean Differences Between the CPT-CA and CPT-MA Scores and Results of the *t* Tests of the Significance of the Mean Differences*

Colors	Pretty Pyramids		Ugly Pyramids	
	Males	Females	Males	Females
	Mean Diff.	Mean Diff.	Mean Diff.	Mean Diff.
Red	.24	.66	.48	1.31
Orange	1.00	.97	.34	1.14
Yellow	.21	.83	.21	.31
Green	— .07	.69	.45	.52
Blue	.41	.76	.03	.83
Purple	.55	1.62	.23*	.32
Brown	1.55	7.21**	.45	.76
White	6.31**	.14	.24	.34
Gray	— .24	.17	.34	1.10
Black	1.10	.41	.41	.83
	.10	.69	.00	.90
	.59	3.47**	.00	1.77

*Negative signs indicate shifts of means in the opposite of the expected direction.

*Significant at or beyond the 5% level of confidence.

**Significant at or beyond the 1% level of confidence.

TABLE III—Means and Standard Deviations for the CPT-MA Scores of the Mentally Defective Ss and Results of the *t* Tests of Differences of the Means with Those of the Norm Group of Normal Ss

Colors	Pretty Pyramids		Ugly Pyramids	
	Males	Females	Males	Females
	Mean	S.D.	Mean	S.D.
Red	5.28	2.23	6.31	1.42
Orange	5.00	2.30	6.70	1.80
Yellow	4.86	2.28	6.17	1.36
Green	5.06	2.00	5.52	1.50
Blue	4.97	1.99	6.83	1.67
Purple	6.59	1.86	6.83	1.67
Brown	5.55	2.35	6.31	1.51
White	5.83	1.65	6.32	1.88
Gray	1.97	2.25	6.11	1.65
Black	5.55	2.01	4.79	1.41
			1.66	1.42

*Significant at or beyond the 5% level of confidence.

**Significant at or beyond the 1% level of confidence.

the Pretty Pyramids for Blue and Brown are related to the Ss' intellectual retardation. This holds true for Purple also, but since a significant difference from the norm group remains after the MA correction, it must be assumed that the original interpretation of emotional disturbance is supported. The differences reported for the female Ss on the Pretty Pyramids for Orange, Blue, and Purple may also be accounted for by the Ss intellectual retardation.

For the Ugly Pyramids, rather different results obtain. The differences from the norm group for male Ss on White and Gray disappear after the MA correction and may thus be accounted for by intellectual retardation alone. All other significant differences remain and thus reflect other personality differences. Intellectual retardation, however, is involved to some degree in the differences for Red, Blue, and Brown for the male Ss and in the differences for Yellow and White for the female Ss, as evidenced by the fact that the MA correction significantly reduces the magnitude of differences without reducing them to insignificance.

It should be noted that the hypotheses used for the *t* tests assessing the effect of the MA correction do not protect against the alternative that the re-scoring procedure might result in some CPT-MA score means deviating further from the norm-group means than was true for the CPT-CA scores. This alternative in fact does occur for all those variables for which differences have negative signs in Table II. The logical explanation for these shifts is that the performance of mental defectives on such variables appears closer to the norm than is actually the case, with intellectual defect serving as a suppressor variable. Since the appropriate *t* tests for this alternative use hypotheses which are the converse of the ones stated above, the values of *t* with negative signs in Table II may simply be evaluated with respect to their significance for the alternate case.

Inspection of Table III shows that suppressor effect results in significant differences from the norm group for the female Ss alone and is restricted to Yellow on the Pretty Pyramids and Red, Orange, Purple, and Brown on the Ugly Pyramids. Intellectual defect as a suppressor variable also increases the magnitude of significant differences with the norm group after MA correction for male Ss for Ugly Orange and for the female Ss for Ugly Black.

DISCUSSION

The above results clearly show that mentally defective children deviate from the CPT norms both on scoring variables related to their intellectual defect as well as on others which may be related to personality deviations concomitant with but not related to such deficit. It now remains to determine whether data on the discrepancies due to mental defect can be used to aid classification.

Test records were pulled at random from our files of normal children matched by age to the defective Ss to assess the effectiveness of classification criteria derived from the results of the present study. It may be seen from a comparison of Tables I and III that, for both boys and girls, at least two scoring variables may be related to intellectual deficit, with mean scores showing differences from the norm group significant at the 1% level of confidence. As classification criteria it is now specified that a given S, to be classified as mentally defective, should score above the normal mean on the variable on which the defective group was high and should score below the mean on the variable on which the defectives were low. The appropriate criterion variables turn out to be low blue and high brown for boys and high orange and purple for girls. Table IV shows the results of classifying the defectives and matched normals and indicates that this procedure has been able to classify correctly about 78% of the 116 records involved. Cross-validation with independent

samples is required, of course, before this procedure can be recommended as an useful adjunct to the diagnosis of mental deficiency.

TABLE IV — Classification of Mental Defective/Normal by Means of CPT Scoring Variables

Males			
Criteria: Blue, Sten 5 and below; Brown, Sten 6 and above			
Predicted Classification			
Observed Classification	Normal	Defective	
Normal	24	5	$r_{tot} .80$
Defective	7	22	
Females			
Criteria: Orange, Sten 6 and above; Purple, Sten 6 and above			
Predicted Classification			
Observed Classification	Normal	Defective	
Normal	25	4	$r_{tot} .76$
Defective	10	19	

What about the meaning of the discrepancies between defectives and normals on those variables which could not be related to defects of intelligence? It has already been indicated that the component of the difference between defectives and normals for Pretty Purple which could not be accounted for by the MA correction may be taken as an indication of emotional disturbance. This makes good sense since evidence abounds that, at least among the mildly defective, emotional problems are frequently concomitant with the intellectual deficit. The discrepancies between defectives and normals on the Ugly Pyramids will generate further hypotheses about the more typical deviant adjustment patterns likely to be prevalent in the defective group. In contrast to the Pretty Pyramids, high scores here are said to reflect modes of affect expression which are either consciously rejected or are latent; or alternate modes which fail to find current expression because of their low order in the S's response hierarchy (Heiss, Honsberg and Karl, 1955).

Proceeding from the interpretive ra-

tionale suggested by Heiss (1952), the high Ugly Red may reflect denial of explosive uncontrolled aggressiveness which may be latent in the defective boys but is not permitted open expression due to the institutional control. The high Ugly Orange similarly may reflect denial of needs for externalization of affect as related to interpersonal function and the Ugly Yellow to denial of needs for emotional investment in object relationships. In the defective girls, the last relationship also applies, but they do not show discrepancies from the normals on Ugly Red or Orange. This suggests either that their needs for aggression and externalization may be lower than for the boys, or may find some other more acceptable outlet. The high Ugly Blue for both boys and girls can probably be attributed to rejection of rationalization as a mode of affect control. The high Ugly White for the girls may suggest schizoid withdrawal tendencies as a denied alternate mode of affect control. This may be the corollary of denied aggressivity as expressed by the high Ugly Red in the boys.

Low scores on the Ugly Pyramids may either be an affirmation of the stability of preferred response modes expressed by corresponding high scores on the Pretty Pyramids, or they may suggest consciously preferred alternate response modes. Thus, the low Ugly Brown in the boys may reflect tendencies towards stubbornness and general resistance to being dominated by others. The low Ugly Black for both boys and girls suggest that depressive mood controls in terms of flattening of affect may be an alternate mechanism. It should be kept in mind that these statements are merely intuitive interpretations of the data and will need to be validated by further studies.

Some comment is in order also as to the color preference behavior of the defective Ss since the comparisons with the norm group used Sten scores and thus gave information about deviation from the norms but not on absolute color preference. Such order of prefer-

ence may, however, be obtained from the Ss' performance on the Pretty Pyramids. It was found that the first ranked colors were Red, Green, and Blue for the boys and Red, Blue, and Green for the girls. The rank order is the same as that for normal girls, but the order of Blue and Green is reversed for the boys (Schaie, 1961c). These results are consistent with most studies of color preference in normal children, which typically show some permutation of Blue, Red, and Green to be the first three preferences (Garth, 1924; Winch, 1909). It appears then that differences between normals and defectives are reflected first of all in the choice of most and least preferred colors as well as the intensity of choice for a specific color. Thus, if response to color is to be used diagnostically, a more complex analysis such as the one possible with the CPT is required.

SUMMARY

The Color Pyramid Test (CPT) was administered to a group of 29 girls and 29 boys who were patients in a State Home for the Mentally Defective. Their scores were compared with those of a norm group of children in the Public School system. When norm tables are entered at the Ss' chronological age significant differences emerge for most of the color choice scores. It was hypothesized that some of these differences could be the effect of intellectual deficit alone. Scores were therefore re-entered into the norm tables at the Ss' Mental Age as estimated by the 1937 Stanford-Binet test. Several of the discrepancies between the normals and defectives disappear as the result of this correction, but others remain. It is suggested that those variables on which differences disappeared due to the MA correction may be useful as classification criteria and validity data are presented showing that approximately 78% of the Ss could be correct-

ly classified from their CPT records as to their respective membership in the groups of defective and normal Ss. A tentative interpretation of the personality attributes of the defective group as inferred from the differences in CPT means not accounted for by intellectual defect is also offered.

REFERENCES

- Garth, R. T. A color preference scale for 1000 white school children. *J. exp. Psychol.*, 1924, 7, 233-241.
- Guilford, J. P. & Smith, Patricia C. A system of color-preferences. *Amer. J. Psychol.*, 1959, 72, 487-502.
- Heiss, R. Der Farbpseudotest. (The Color Pyramid Test). *Psychol. Rdsch.*, 1952, 3, 1-11.
- Heiss, R. and Hiltmann, H. Der Farbpseudotest nach Max Pfister. (The Color Pyramid Test according to Max Pfister). Bern: Huber, 1951.
- Heiss, R., Honsberg, I., & Karl, H. Vorläufige Mitteilung ueber die Verwendung von "haesslichen Pyramiden" im Farbpseudotest. (Preliminary communication on the use of the "ugly pyramids" in the Color Pyramid Test.) *Z. diagnost. Psychol.*, 1955, 3, 106-124.
- Karl, H. Die Methodik des Mittelprotokolls im Farbpseudotest. (Methodology of the average protocol in the Color Pyramid Test.) *Z. exp. angew. Psychol.*, 1956, 4, 123-145.
- Norman, R. D., & Scott, W. A. Color and affect: a review and semantic evaluation. *J. gen. Psychol.*, 1952, 46, 185-223.
- Schaie, K. W. Scaling the association between colors and mood-tones. *Amer. J. Psychol.*, 1961, 74, 266-273. (a)
- Schaie, K. W. A Q-sort study of color-mood association. *J. proj. Tech.*, 1961, 25, 341-346. (b)
- Schaie, K. W. Child and adolescent norms for the Color Pyramid Test. Unpubl. manuscript, Univ. of Nebraska, 1961. (c)
- Schaie, K. W. Die Validitaet des Farbpseudotests. (The validity of the Color Pyramid Test.) *Diagnostica*, in press, 1962.
- Schaie, K. W. The Color Pyramid Test: A non-verbal technique for personality assessment. *Psychol. Bull.*, in press, 1963.
- Stacey, C. L. and Reynolds, W. F. The color preferences of a group of subnormals. *Tr. Sch. Bull.*, 1953, 50, 59-64.
- Winch, W. H. Color preferences of school children. *Brit. J. Psychol.*, 1909, 3, 42-65.

Received March 5, 1962

The "Meaning" of the Bender-Gestalt Test Designs to Psychiatric Patients

HERBERT C. SCHULBERG¹ AND ALEXANDER TOLOR
Fairfield State Hospital, Newtown, Connecticut

In recent years there has been a growing interest in using the Semantic Differential (Osgood, Suci, & Tannenbaum, 1957) to determine experimentally the connotative meaning of various test stimuli. A major purpose of several studies has been to provide clinicians with empirically derived information about projective stimuli in order to facilitate the interpretation of patients' responses and to reduce the need to rely on unverified hunches. This experimental approach is exemplified by Rabin's (1959), and Zax and Loisel's (1960) studies with the Rorschach inkblots, and Tolor's (1960) investigation with the Bender-Gestalt Test designs.

The earliest experimental study of the stimulus value of the Bender test was that of Suczek and Klopfer (1952) in which college students were instructed to free-associate to each design; various interpretative hypotheses were then inferred from the resulting clusters of associations. The associational approach was also used by Tolor (1957) with a group of neuropsychiatric patients in a study of the structural characteristics of the designs. In both of these studies it was concluded that the Bender figures vary considerably in their stimulus value and that this variance requires further explication. Tolor's (1960) more recent use of the Semantic Differential technique for determining the connotative meanings of the Bender figures in college students represents a more sophisticated attempt to obtain empirical information about the test, and it was again found that the designs vary considerably in the descriptions which they elicit.

The possible importance of this finding becomes apparent when one considers that the quality of the copy performance on the Bender-Gestalt Test tends to differ appreciably for discrete psychiatric subgroups. For example, Robinson (1953) found that the Bender productions of paretics tend to be less accurate than those of schizophrenics, and Lonstein (1954) found that the group performance of schizophrenics was significantly worse than that of neurotics. The results of Bowland and Deabler's (1956) study, along with several others, indicated that the quality of the copy performance can generally be ordered from neurotics to psychotics to organics, with the first group demonstrating the least impairment and the last group showing the greatest degree of impairment in Bender drawing fidelity.

In addition, there is some support for the contention that specific Bender drawing deviations are related to certain adjustment disturbances and to other test measures having dynamic significance. For example, Clawson (1949) found that various types of childhood difficulties, such as withdrawn or acting-out behavior, have their counterparts in certain Bender test manifestations. She also demonstrated significant relationships between selected Rorschach determinants and content categories on the one hand and Bender test performance on the other hand. It is therefore not unlikely that different styles of Bender performance may be characteristic for various diagnostic groups.

The question that now presents itself is whether or not the apparent group differences in visual-motor functioning are a function, at least in part, of the varying connotative meanings attributed to the designs. Although

¹ Now N.I.M.H. Postdoctoral Fellow, Harvard School of Public Health, 55 Shattuck St., Boston, Mass.

the Bender-Gestalt Test may simply be conceived of as a test of maturation, it is frequently regarded as a measure of psychomotor efficiency related to general ego-intactness and as a projective device. When viewed in this latter way, one could easily entertain the possibility that the Bender design associations, if they were relatively specific for different diagnostic subgroups, would reflect varying drives and adaptive styles contributing also to the group differences noted in reproducing the configurations. Thus, a generally inferior performance in copying the designs, or one which differs qualitatively, may be related to the specific and probably relatively more disturbing stimulus value of certain figures as compared to others rather than being simply the result of impaired perceptual-motor functioning. The purpose of this study, therefore, is to determine the nature of the connotative meanings assigned to the designs by various psychiatric subgroups, and to ascertain whether these descriptive rating patterns differentiate the groups.

METHOD

The same 20-scale form of the Semantic Differential which was used previously by Tolor (1960) was administered in the present study. It includes scales representative of the three major factors which have been consistently obtained with different samples of Ss and with a variety of concepts. The *Evaluative Factor* shows high loading on scales 1, 3, 5, 6, 7, 8, 9, 10, 13, and 16; the *Potency Factor* is represented by scales 2, 4, 12, 15, 18, and 20; and the *Activity Factor* is prominently found in scales 11, 14, 17, and 19.

A variation of the usual group administration procedure was employed. Instead of exposing the figures one at a time to the entire group for a specific time interval, the Bender figures were presented to each subject individually in the form of a mimeographed nine-page booklet with each page containing a different figure

along with the 20 Semantic Differential scales. It should be noted that in the mimeographing process some of the designs inadvertently underwent some slight change so that the stimuli used here cannot be considered to be exact replicas of Bender's original figures. In order to obtain a variety of diagnostic categories Ss were drawn from the whole range of the newly admitted male population on the acute treatment service at Fairfield State Hospital. For several months, weekly group testing sessions were held at which time all patients who had been admitted within the past seven day period were asked to rate the designs on the Semantic Differential. It was considered desirable to obtain responses during the early phases of hospitalization in order to control for the effects of treatment, and to obtain the data while patients were presumably still displaying the acute symptoms of their particular disorders.

The Vocabulary subtest of the WAIS was administered as an approximate measure of intelligence during the same session so that intellectual differences among the diagnostic groups could be determined and, if necessary, controlled. This procedure also provided some gross indication of the patient's ability to understand the verbal instructions given with the Semantic Differential. The responses of patients unable to perform satisfactorily on the Vocabulary subtest were eliminated from further consideration. In addition, non-English speaking patients, and patients who were too confused or disoriented to comply with the task, e.g. hallucinating schizophrenics and senile organics, were also excluded from the study.

The sample consisted of 15 neurotics, 41 functional psychotics, 15 acute organic psychotics, and 45 character disorders (84% of whom were alcoholics). The diagnoses were assigned by the psychiatric staff either after several weeks of evaluation or at the time of discharge. All patients were male; the mean age of the various

TABLE I—Mean Semantic Differential Ratings Assigned by the Combined Group of Psychiatric Patients (N = 116) to the Nine Bender Designs

	Design A	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8
1. good-bad	2.99	3.28	2.47	3.35	3.85	3.46	4.14	3.91	2.19
2. large-small	3.98	5.04	4.32	4.37	3.63	4.11	2.76	3.93	3.14
3. beautiful-ugly	3.68	4.27	3.26	3.87	4.37	3.93	4.46	1.13	2.93
4. strong-weak	3.97	3.86	3.41	3.75	4.20	3.87	4.00	3.88	2.90
5. clean-dirty	2.33	2.84	2.36	2.98	3.20	3.01	3.76	3.06	2.11
6. tasty-distasteful	4.45	4.21	3.58	4.00	4.51	4.10	1.51	1.43	3.51
7. valuable-worthless	3.98	4.00	3.72	3.82	4.26	4.00	1.16	4.21	3.01
8. kind-cruel	3.15	3.62	3.17	3.41	3.58	3.50	3.91	3.74	3.15
9. pleasant-unpleasant	3.46	3.54	3.04	3.36	4.00	3.56	4.15	1.01	2.77
10. happy-sad	4.98	3.81	3.10	3.45	3.65	3.49	3.88	1.03	3.13
11. ferocious-peaceful	4.42	4.96	4.94	4.58	4.30	1.77	4.03	1.20	1.80
12. heavy-light	2.71	4.53	4.36	4.02	3.81	4.05	3.95	3.78	3.91
13. clear-hazy	4.39	2.90	2.38	2.81	3.07	3.07	3.15	4.23	2.56
14. hot-cold	4.49	4.43	4.34	4.32	4.25	4.15	1.80	4.15	1.28
15. thick-thin	2.88	4.58	4.09	4.35	4.01	1.20	4.23	3.65	3.71
16. honest-dishonest	3.63	2.97	2.87	3.29	3.10	3.37	3.75	3.61	2.90
17. active-passive	4.29	3.69	3.58	3.54	3.95	3.74	3.71	3.68	3.36
18. rough-smooth	3.78	4.65	5.11	4.38	4.19	3.95	3.78	3.65	4.47
19. fast-slow	4.15	3.51	3.79	3.59	4.07	3.86	3.85	3.14	3.64
20. rugged-delicate		4.41	4.53	1.05	1.07	4.16	3.92	3.87	1.11

groups ranged from 36.5 years to 43.3 years, with none of the age differences being significant. The patients' estimated intelligence, based on the mean vocabulary score, fell in the Average range. The vocabulary items had been scored independently by two judges who attained an inter-rater reliability of .96. The mean educational level attained was 10.5 grades.

RESULTS

Preliminary to the comparison of the four diagnostic groups' Semantic Differential ratings, an attempt was made to determine the effect of intelligence. Accordingly, the WAIS Vocabulary scores of the four diagnostic subgroups were compared with each other by six *t* tests. It was found that the mean scores of 35.86, 42.97, 36.53, and 43.83 for neurotics, functional psychotics, acute organic psychotics, and character disorders, respectively, were not significantly different. It was therefore concluded that these patients were drawn from the same population in regard to intelligence.

Patient Differences

The primary data of this study consist of the 180 choices (20 scales x 9 Bender designs) made by each subject. Since the mean rating for each group on every scale was compared with that of every other group, a total of 1,080 comparisons (6 possible group combinations x 180 selections) were computed.² Of 1,080 *t* tests, only 47 were found to be significant at the .05 level or beyond, a number that is very likely due to chance. Moreover, there were no consistent trends in regard to the specific significant pairs of groups or the specific designs which yielded these small number of significant differences. It would appear, therefore, that there are no essential differences in the connotative meanings which the Bender designs evoke via the Semantic Differential as related to diagnostic classification.

² The authors wish to express their appreciation to Mrs. Myra Tolchin for her assistance in the statistical analysis of the data.

Design Differences

In view of the high degree of uniformity of the Bender design characterizations, it is possible to generalize about the connotative meanings assigned by psychiatric patients to each of the figures. In Table I the lower the numerical score, the more the patients felt the left-hand adjectives to be descriptive of the design. Conversely, higher scores are associated with the adjectives found on the right-hand side of the scales. The ratings theoretically can range from a low of 1 to a high of 7, reflective of the seven points that may be selected between each bipolar scale.

An attempt was also made to determine whether the design variance, as determined by ratings on the 20 scales, is significant when all subgroups are combined. One-way analyses of variance, computed for every scale, revealed significant differences at the .05 level or beyond for the designs on fourteen of the twenty pairs of adjectives. The only scales on which the designs could not be significantly differentiated by the patients were scales 12, 14, 16, 17, 19, and 20. This finding indicates that, in general, patients respond differentially to the nine Bender configurations in regard to the

TABLE II — Analysis of Variance for Each Semantic Differential Scale for All Patients Combined

Scale	F	P
1	8.50	.01
2	20.80	.01
3	7.00	.01
4	5.25	.01
5	9.13	.01
6	4.75	.01
7	2.70	.05
8	2.67	.05
9	7.25	.01
10	7.83	.01
11	3.87	.01
12	1.33	N.S.
13	3.08	.05
14	.63	N.S.
15	4.60	.01
16	2.09	N.S.
17	.80	N.S.
18	4.14	.01
19	2.25	N.S.
20	1.21	N.S.

connotative meanings assigned to them. Table II presents the results of the analysis of variance for each scale when all patient groups are combined.

Factorial Scores

Another approach to the problem of whether the Bender designs have differential meanings for various psychiatric subgroups involves the determination of responses to the Evaluative, Potency, and Activity factors which are reflected in the scales of the Semantic Differential. An analysis of variance indicated that there were no significant between-patient group differences. Thus, there appears to be no substantial relationship between the representation of any specific factor and diagnostic category.

Response Sets

In view of the susceptibility of rating scales, such as the Semantic Differential, to response sets, especially when used with psychiatric patients, the investigators attempted to determine whether any of the subgroups made significantly greater use of the neutral category, i.e., the mid-point, on the seven-point scale. The incidence of the use of the mid-point as compared to the remaining spaces was determined for all of the scales and for each design, and chi-square comparisons were made for the diagnostic subgroups. The results indicated that of 1,080 chi-squares, only 17 were significant, a number that does not exceed chance expectations. The data there-

fore do not indicate the greater prevalence of the neutral category response set in any of the psychiatric subgroups as compared to others.

Another way in which possible differential response sets may be subjected to scrutiny is by a comparison of the relative spread of responses on the Semantic Differential. Here, the focus of attention is on the frequency with which each of the seven points on the scales are used by different groups of patients. When the choices for all of the nine Bender-configurations are treated in combination, highly significant chi-squares are obtained for each of the six between-group comparisons, indicating that responses are differentially dispersed throughout the seven-point scales for the different diagnostic samples. Further analysis of these significant differences revealed that they were derived chiefly from the functional psychotic group's excessive reliance on one of the extreme categories on the seven-point scales. Table III presents information regarding the differences in response variability for the four diagnostic groups.

DISCUSSION

In view of the fact that various diagnostic groups appear to differ in the ways in which they reproduce the nine geometric designs on the Bender-Gestalt Test, a critical need was seen for studies elucidating the stimulus value of the figures to various psychiatric subgroups. An adequate comprehension of the symbolic, affective, and as-

TABLE III—Dispersion of Ratings on the Semantic Differential for Each Diagnostic Group

Diagnostic Groups	Frequencies							Total Ratings
	1	2	3	4	5	6	7	
Neurotics (N=15)	297	263	388	1070	270	144	267	2699
Functional Psychotics (N=41)	1877	392	631	2418	526	452	1060	7356
Acute Organic Psychotics (N=15)	421	220	290	971	213	195	389	2699
Character Disorders (N=45)	1300	719	854	2822	763	740	839	8037

$$X^2 = 658.91$$

$$d.f. = 18$$

$$p = < .001$$

sociational impact which each of the Bender designs produces is thus regarded to represent an important intermediate stage in the final evaluation of the dynamic significance of the drawing deviations in patients. To this end the authors used a method which combines controlled association and scaling procedures to differentiate the connotative meanings assigned to each of the designs by relatively homogeneous groups of hospitalized psychiatric patients.

The general failure of this study to obtain major differences in the meanings attributed to the designs by the four diagnostic groups used here suggests that individuals who are so incapacitated as to require in-patient treatment do not perceive the Bender stimuli differentially. Whether the general severity of the disturbances represented in these samples obscured differences which the inclusion of non-hospitalized patients and normals might have revealed requires further investigation. In addition, we are not certain whether the samples employed in this study would have differed significantly in their drawing reproductions of the configurations. In the event that group differences in the drawings could not be discerned, then the present findings could possibly be accounted for on the basis of generally equivalent ego-strengths amongst these groups of patients. On the other hand, in the event that the copy productions were to differ appreciably among groups, one would have to conclude that the meanings assigned to the stimuli by patients are unrelated to test performance.

This study confirmed previous findings that there are design differences in regard to connotative meanings elicited, and extended these findings to psychiatric patients. In general, Design A may be characterized as good, clean, peaceful, clear, and honest. Design 1 was described as small, clean, peaceful, clear, and honest. The most common ratings for Design 2 were good, clean, pleasant, peaceful, clear,

honest, and smooth. Design 3 is clean and clear; Design 6 is described as large, while Design 8 is regarded as good, beautiful, strong, clean, valuable, pleasant, clear, and honest. It will be noted that Designs 4, 5, and 7 tend to be rated neutral on all the scales so that no specific characterization is warranted.

A further observation based on the listing of descriptive terms pertains to the general disinclination of patients to select that polar end of the continuum that would express an unfavorable evaluation, a low potency, or a low activity. This result stands in sharp contrast to Tolor's (1960) previous findings with college students in which case there was almost equal representation of both ends of the scale continua. It thus seems that a powerful social conformity set is operating with psychiatric patients.

In regard to the six scales on the Semantic Differential which failed to differentiate between Bender designs, all of them appear toward the latter half of the series of twenty scales. One is therefore tempted to speculate whether the failure of these pairs of adjectives to distinguish between designs was a function of the increased fatigue and accompanying disinterest in the patients as they progressed from the first to the last scales in the series. It is likely that patients are much more subject to these factors which attenuate any differences between the polar ends of the scales than normals would be.

Unlike Zax, Loiselle and Karras (1960) who found group differences between schizophrenics and hospitalized medical patients, on the one hand, and college students, on the other hand, in tendency to use the neutral category in their ratings, the present authors found no general group differences in this respect for their four psychiatric samples. The lack of group differences in the present study may very well be due to the present groups' similarities in age and intelligence which are factors that

seem to determine the ability to make discriminations in ratings. However, another response set, namely, the significantly greater reliance of the functional psychotic group on the extreme left-hand category of the seven-point scale was noted. This behavior suggests that psychotics tend to make the first response available in complex situations.

SUMMARY

In view of the differing abilities of psychiatric subgroups to copy the Bender designs, it was sought to determine whether group variations in performance could be associated with differences in the connotative meanings attributed to the designs. Ratings of each Gestalt on twenty scales of the Semantic Differential were obtained from groups of neurotics, functional psychotics, character disorders, and acute organic psychotics. It was found that all subgroups described the designs in a basically similar fashion. However, patients responded differentially to the nine Bender figures in regard to the connotative meanings assigned to them. No significant differences were noted in the patients' use of the Semantic Differential's Evaluation, Potency, and Activity factors. A significantly greater use of one extreme category of the seven-point scale was found for the psychotics.

REFERENCES

- Bowland, J. A., & Deabler, H. L. A Bender-Gestalt diagnostic validity study. *J. clin. Psychol.*, 1956, 12, 82-84.
- Clawson, Aileen. The Bender Visual Motor Gestalt Test as an index of emotional disturbance in children. *J. proj. Tech.*, 1959, 27, 198-206.
- Lonstein, M. A validation of a Bender-Gestalt scoring system. *J. consult. Psychol.*, 1954, 18, 377-379.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.
- Rabin, A. I. A contribution to the "meaning" of Rorschach's inkblots via the semantic differential. *J. consult. Psychol.*, 1959, 27, 368-372.
- Robinson, Nancy M. Bender-Gestalt performances of schizophrenics and parotics. *J. clin. Psychol.*, 1953, 9, 291-293.
- Suzek, R. F., & Klopfer, W. G. Interpretation of the Bender-Gestalt Test: the associative value of the figures. *Amer. J. Orthopsychiat.*, 1952, 22, 62-75.
- Tolor, A. Structural properties of the Bender-Gestalt Test associations. *J. clin. Psychol.*, 1957, 13, 176-178.
- Tolor, A. The "meaning" of the Bender-Gestalt Test designs: a study in the use of the semantic differential. *J. proj. Tech.*, 1960, 24, 433-438.
- Zax, M., Loiselle, R. H., & Karras, A. Stimulus characteristics of Rorschach inkblots as perceived by a schizophrenic sample. *J. proj. Tech.*, 1960, 24, 439-443.
- Zax, M., & Loiselle, R. H. Stimulus value of Rorschach inkblots as measured by the semantic differential. *J. clin. Psychol.*, 1960, 16, 160-163.

Received March 26, 1962

A Scoring Rationale for the Sargent Test of Insight into Human Motives

ELLIOTT P. SCHUMAN
Long Island University

The Test of Insight into Human Motives by Sargent (1953) appears to have many of the advantages of a projective test (encouraging, for example, enough free expression and fantasy to permit the revelation of ordinarily more guarded thoughts and feelings), and also those advantages which would logically accrue to: (a) a paper-and-pencil test (facilitating research); (b) a test ostensibly concerned with human motivations (which would probably be more interesting to Ss and might later lend itself to group discussions of the problem areas treated); (c) a test eliciting reactions to more direct questions concerning the actions and motivations of the central character. The test elicits responses to social problem situations from which measures of both content and structure can be obtained. Four 15-item (each item is referred to by Sargent as an "armature") forms are provided, two for males and two for females, with one of each pair apparently more suitable for younger, and the other for older, persons. In addition, Sargent provided for short forms by starring ten of the items in each form. While recommending the 15-item form for individual case study and recognizing the administrative advantages of the 10-item form for groups, she conceded that "any number of armatures may be used" (1953, p.18). All items refer to another person of the same sex and approximate age. To each item the S is asked: (a) What did he do and why? (b) How did he feel? One representative item is: A young man gets a good deal of razzing because he spends his week-ends at home instead of dating.

Sargent's scoring procedures deal with three principal dimensions—affect, defense, and malignancy (by which she means deviant verbaliza-

tions). The writer (1958) dealt with four of Sargent's twelve affect subscores (interpretation of situations as frustrating, interpretation of situations as facilitating, positive perception of people, and negative perception of people), all three of her defense subscores (elaboration, evaluation, and qualification), and all six of her malignancy subscores (irrelevant feeling, bizarre productions, use of the first person, a solution based upon elaboration of the item, no solution, and no feeling).

Adequate reliability appeared vital to fruitful utilization of the test. Sargent's reliability data, reported primarily in terms of percentage of agreement, do not provide too clear a picture of the extent to which different scorers will agree, the same scorer will agree with scores previously assigned by him, or Ss on a later administration will reproduce scores earned on a previous administration. When the writer and another psychologist used the scoring rules provided by Sargent to score protocols supplied to illustrate her scoring procedures, it became apparent that a considerably more refined scoring rationale would have to be developed. More comprehensive and detailed rules were then formulated. An administration of Form I-Men, the form apparently more appropriate to the experimental sample, to an 11-subject pilot group indicated the desirability of: (a) shortening the 15-question form; (b) clarifying an ambiguity in a question referring to one of the problem situations; (c) changing certain portions of the instructions to make them more applicable to the conditions of the actual administration; (d) insuring the inclusion within the response, not only of the action proposed and the reason therefor, but also of the feeling com-

ponent; (e) insuring responses to all problem situations. Accordingly, the following changes were made: (a) the standardized 10-question form was employed; (b) question number four was reworded for greater clarity; (c) the instruction to "label parts (a) and (b) separately" was added; (d) the instruction "You may write for an hour. Write first on the items that interest you most, since you may not have time to answer all the questions in the time allowed" (Sargent, 1953, p.260) was deleted. The form of the Insight Test actually employed is as follows:

TEST OF INSIGHT INTO HUMAN MOTIVES

Instructions

Insight into other people helps us to get along with them. This is a test of your ability to "see into" others. This requires both imagination and the ability to "put yourself in someone else's place."

On the following pages you will find a number of situations described very briefly. After each you will find two questions which you are asked to answer in writing on the blank pages provided. Please number your answers to correspond to the questions on which you are writing, and label parts (a) and (b) separately.

Notice that the persons involved in the situations are not described. This allows you to use your imagination as to what sort of people the characters might be.

There are no right and wrong answers, but your explanations should show understanding of the characters as you see them.

1. A young man is working or studying away from home. He gets a letter from his mother, after the death of his father, asking him to move back home.
 - a. What did he do and why?
 - b. How did he feel?
2. A husband is able to support his wife. But she tells him she wants to keep on working because she prefers to be financially independent.
 - a. What did he do and why?
 - b. How did he feel?
3. A young man gets a good deal of razzing because he spends his weekends at home instead of dating.
 - a. What did he do and why?
 - b. How did he feel?
4. A young man's father has always looked

forward to having his son take over his business and has educated him for it. The son becomes interested in another vocation.

- a. What did the son do and why?
- b. How did he feel?
5. A young man discovers that a girl to whom he is engaged has had a very bad reputation in the past.
 - a. What did he do and why?
 - b. How did he feel?
6. A young man is caught in petty theft. His employer offers to let him keep his job if his parents will vouch for him. Otherwise he will be forced to look for another job without references.
 - a. What did he do and why?
 - b. How did he feel?
7. A young man has a date with a girl he has been seeing regularly. At the last minute she breaks it without explanation.
 - a. What did he do and why?
 - b. How did he feel?
8. A young man promised a friend to help him with some work. Another friend asks him to go on a blind date with a girl who is said to be very attractive.
 - a. What did he do and why?
 - b. How did he feel?
9. A young man gets the impression that others are discussing him. On several occasions he thinks conversation has stopped or the subject changed when he entered the room.
 - a. What did he do and why?
 - b. How did he feel?
10. A young man has been feeling quite ill but the doctor tells him he is not seriously sick and it will do him no harm to go on leading a normal life.
 - a. What did he do and why?
 - b. How did he feel?

Scoring of the pilot study protocols by three judges revealed the expediency of: (a) condensing the rules; (b) simplifying one of the subcategories; (c) presenting a greater number of recommendations for scoring individual words and phrases, particularly for those responses concerned with situations and people. The rules were revised with the foregoing objectives in mind. The investigator and one non-clinically oriented graduate student served as the two judges for the study proper. It was felt that the in-

terjudge reliability would be tested more rigorously if clinical acumen were not required. After the second judge had been given an opportunity to study the rules and to discuss the intricacies of the scoring rationale, six sample protocols were scored independently by the two judges, and the scoring compared and discussed. The discussions which preceded and followed the scoring of the sample protocols permitted modifications of the rules which served further to sharpen the scoring rationale. Thus, Sargent's rules were revised on three separate occasions prior to scoring of the protocols (both test and retest) of the experimental sample.¹

The Ss' responses, transcribed into typescripts, were scored independently by the two judges for 13 subcategories comprising four categories:

1. Defense, *D*, the criterion of defensive verbalizations, the sum of:

- a. Qualification, *Q*;
- b. Evaluation, *Ev*;
- c. Elaboration, *El*.

2. Deviant Verbalization, *M* (Sargent's Malignancy category, a purported indicator of pathology), the criterion of deviant verbalizations, the sum of:

- a. No Solution, *OSol*;
- b. No feeling, *OA*;
- c. Personal Pronoun, *PP*;
- d. Subjectivism, *Sj*;
- e. Irrelevant Feeling, *Ir*;
- f. Unreal Solution, *Unr*.

3. Direction of Situational Appraisal, *FA - FR*, the criterion of interpretation of situations as frustrating, the difference of:

- a. Situations as Facilitating, *FA*;
- b. Situations as Frustrating, *FR*.

4. Direction of Personal Appraisal, *P - N*, the criterion of negative perceptions of people, the difference of:

- a. People as Positive, *P*;
- b. People as Negative, *N*.

¹The 16 pages of rules which emerged have been deposited with the American Documentation Institute. Order Document No. 7291, remitting \$1.75 for 35-mm microfilm or \$2.50 for 6 by 8 in. photocopies.

Single tallies were made for each of the 13 subcategories on the basis of concepts, sentences, phrases, or specific words, each judge's subcategory score for a particular subject comprising the total number of tallies assigned to the particular protocol. The final subcategory score for each *S* was obtained by summing the scores assigned by each of the two judges.

To take advantage of judgments based upon a constant set, and to minimize changes in scoring associated with the greater familiarity with the rules which might obtain as more protocols were scored, the following procedure was adopted: (a) the responses to question number one were scored before proceeding to the scoring of the responses to question number two, this sequential order of scoring responses to questions being maintained throughout; (b) the order for scoring protocols consisted, first, of the subsample of 16 second administrations in alphabetical order, and second, of the sample of 80 first administrations in alphabetical order; (c) scoring of the responses to question number one was begun with the first protocol, scoring of the responses to question number two with the eleventh, scoring of the responses to question number three with the twenty-first, etc.

In order to illustrate actual scoring by the two judges, a protocol was selected by choosing the first test protocol alphabetically to include at least one score in each of the subcategories which were scored with sufficient frequency to be included in the correlational matrix. Scores assigned by Judge T are indicated above those assigned by Judge S. The protocol is as follows:

- 1a. I would say that the boy would move back home for a time only.

His mother's immediate need is for con-

El

Ev

solation and companionship. When she becomes acclimated to a life alone, the

Ev

El

boy will be able to resume his studies.

b. The young man, although probably lov-
 ing his parents deeply, will no doubt

feel resentment. It will not be an easy
 thing to leave the place where he has
 sunk in his roots and begin a new life.

Furthermore, this moving may mean a
 change in the boy's vocational or future
 vocational life, thus heightening the re-
 sentment.

2a. He probably would have made her stop
 working as he feels his reason for doing
 so comes ahead of her desire to be finan-

cially independent.

b. A man, when he assumes the position as
 head of a family, should feel that they
 cannot get along without him. This re-

sults in a happier household as either a
 husband or wife must be willing to
 subordinate himself to the other or con-

flict will occur. This attitude makes for

confidence and satisfaction in the hus-

band while the wife can derive no less
 satisfaction from her position as a lov-

ing wife. Therefore, the man feeling his
 confidence and assuredness would be

undermined, to the detriment of his

whole family, would have asked the wife
 to stop working.

3a. This is a case of personalities and if the
 man could stand the razzing and actual-
 ly preferred to stay at home, he would

do so. In any case, he probably would
 have stayed home, as he is frightened to
 take a step in the direction of female

associations.

b. He hates himself for being so afraid of
 girls but nevertheless refuses to date

them. He makes up excuses, both to
 himself and others.

4a. The son would have gone into his own
 choice of profession. In this particular
 case, the love for his father would have
 to be exceedingly great in order for him

to follow his wishes.

b. He feels that he would be ruining his

own life if he went into his father's
 business. His father has lived his life
 and now he should give him a chance

to live his.

5a. If the young man respects the opinions
 of society more than his love for the
 girl he will break the engagement. How-

ever, if the girl's love for him is genuine
 and vice versa and they can make a start
 some place else, then he will maintain

his relationship with her.

b. Although his love for her may be great,
 he feels that her reputation will hold

back any advancements he may make in
 life. Thus his social conscience will pre-

vent him from marrying her.

6a. He would leave his present job and try
 for a new start, however hard it may be.
 He cannot be happy or respected in his
 present position with the backward

glances people will throw at him. He
 will tend to become irritable and

touchy. Thus with a new job he can

start with no strikes against him and although it may be tough, it will be

Ev
Ev

worth it.

- b. He feels that he will be too dependent on other people and his individualism

Q
Q

will be suppressed if he accepts "charity" and continues with his present position. A new start will allow him to re-

Ev OA
Ev OA

tain this individualism.

- 7a. He would not seek any explanation and pass it off till she could explain it. He does not wish to judge her too hastily and have their relationship broken off

OSol

because he was suspicious.

FR
FR

- b. He feels hurt but is willing to let it go as his feeling for her is dominant over his feeling of hurt.

- 8a. As friendship is worth more to him than the blind date, he will help his

P
P

friend. He does not wish to sacrifice his

Ev

friendship for something so fleeting as a date with an attractive girl.

- b. He feels a little longing for the date he

FR
FR

could have had but sees the benefits of

Ev

a lasting friendship and is readily con-

FA
FA

soled.

- 9a. He would speak to one of the people whom he thought was talking about

El

him and find out just what was wrong.

Q
Q

If he knew what was wrong he would be able to correct it.

Ev FR
Ev FR

- b. He certainly felt self-conscious and no

Ev
Ev

doubt he felt he was being either ridi-

FR

culed or maligned.

Ev
Ev

- 10a. He was no doubt imagining his ills be-

Ev
El

cause of some psychological problem.

El

Since that problem is not eliminated he

Ev

will go on pretending sickness as a de-

Ev
Ev

fense against that problem.

Q
Q

- b. Probably the young man seriously felt he was sick but the doctor tells him something to the contrary and he feels

FR
FR

a letdown but rationalizes the facts with the doctor that he doesn't know what

Q

he is talking about or some such expla-

OA
OA

nation.

Because it seemed that the total number of instances in which a subject directed his attention to situations (or to persons) *per se* would be a datum worthy of investigation, the Situational Appraisal and Personal Appraisal subcategories were summed without regard to direction of appraisal, producing two additional categories. Thus, from the 13 scorable subcategories there emerged six principal categories—Defense, *D*, Deviant Verbalization, *M*, Situational Appraisal, *FA* + *FR*, Direction of Situational Appraisal, *FA* - *FR*, Personal Appraisal, *P* + *N*, and Direction of Personal Appraisal, *P* - *N*.

It seemed prudent to examine the interrelationships between subcategories as well as between principal categories, excluding those which it appeared could not be scored with sufficient reliability, or which were not scored with sufficient frequency. Four of the six Deviant Verbalization subcategories (Personal Pronoun, *PP*, Subjectivism, *Sj*, Irrelevant Feeling, *Ir*, and Unreal Solution, *Unr*) were excluded (primarily because they were scored so infrequently), leaving a total of 15 Insight Test variables based upon judgments—three Defense sub-

categories (*Q*, *Ev*, and *El*), two Deviant Verbalization subcategories (*OSol* and *O1*), two Situational Appraisal subcategories (*FA* and *FR*), two Personal Appraisal subcategories (*P* and *N*), one Defense category (*D*), one Deviant Verbalization category (*M*), two Situational Appraisal categories (*FA + FR* and *FA - FR*), and two Personal Appraisal categories (*P + N* and *P - N*).

Table I presents, for the 15 judged categories of the Sargent Test, the in-

dividual judges' means and standard deviations, and the interjudge coefficients of reliability. Table II presents, for the 15 judged categories and also for amount of writing, the means and standard deviations, and the test-retest coefficients of reliability for a subgroup retested after a period of nine weeks.

The data in Table I demonstrate that the judges interpreted rather similarly the rules adopted for scoring the Sargent Test. Inter-judge reliabilities

TABLE I—Individual Judges' Means and Standard Deviations, and Interjudge Coefficients of Reliability, for Judged Categories of the Sargent Test

	Judge S		Judge T		Interjudge r
	\bar{X}	s	\bar{X}	s	
<i>Q</i> (Qualification)	12.65	10.91	12.49	10.64	.99**
<i>Ev</i> (Evaluation)	9.61	6.07	8.28	5.97	.94**
<i>El</i> (Elaboration) ^a	1.55	4.18	2.70	2.89	.81**
<i>D</i> (Total defense)	26.81	13.57	23.46	12.53	.96**
<i>OSol</i> (No solution)	2.78	2.54	2.99	2.58	.96**
<i>OA</i> (No feeling)	4.14	2.53	4.50	2.66	.96**
<i>M</i> (Total deviant verbalization)	7.04	3.99	7.59	1.18	.96**
<i>FA</i> (Situations as facilitating)	1.50	1.31	1.26	1.22	.88**
<i>FR</i> (Situations as frustrating)	6.74	4.70	6.13	3.96	.96**
<i>FA-FR</i> (Difference plus 50)	44.76	4.82	45.14	4.18	.97**
<i>FA+FR</i> (Sum)	8.24	4.94	7.39	4.11	.94**
<i>P</i> (People as positive)	2.11	1.73	1.80	1.64	.84**
<i>N</i> (People as negative)	1.89	2.22	1.46	1.70	.94**
<i>P-N</i> (Difference plus 20)	20.23	2.71	20.34	2.28	.88**
<i>P+N</i> (Sum)	4.00	2.91	3.26	2.44	.89**

N = 80

**Significant at .01 level.

TABLE II—Means and Standard Deviations on Test and Retest, and Test-Retest Coefficients of Reliability

	Test		Retest		Test-Retest r
	\bar{X}	s	\bar{X}	s	
Amt. (Amount of writing)	41.57	22.37	36.43	16.81	.92**
<i>Q</i> (Qualification)	19.94	14.18	17.44	13.23	.44*
<i>Ev</i> (Evaluation)	21.88	12.21	18.25	14.52	.63**
<i>El</i> (Elaboration)	8.94	9.28	4.75	3.01	.67**
<i>D</i> (Total defense)	50.75	22.51	40.44	22.43	.63**
<i>OSol</i> (No solution)	5.00	4.44	5.56	5.41	.59**
<i>OA</i> (No feeling)	5.94	4.64	8.31	4.51	.49*
<i>M</i> (Total deviant verbalization)	11.13	7.85	13.88	8.56	.59**
<i>FA</i> (Situations as facilitating)	2.88	2.39	2.38	3.43	.56*
<i>FR</i> (Situations as frustrating)	14.81	8.67	12.31	7.11	.51*
<i>FA-FR</i> (Difference plus 50)	38.06	9.49	40.06	8.84	.55*
<i>FA+FR</i> (Sum)	17.69	8.46	14.69	6.80	.46*
<i>P</i> (People as positive)	4.69	2.93	3.88	3.51	.66**
<i>N</i> (People as negative)	4.25	4.60	4.44	4.88	.76**
<i>P-N</i> (Difference plus 20)	20.44	5.03	19.44	6.03	.76**
<i>P+N</i> (Sum)	8.94	5.86	8.31	5.99	.70**

N = 16

*Significant at .05 level.

**Significant at .01 level.

ranged from .81 to .99 with only two, both of which are subcategories, below .88; each of the two principal dimensions, total defense and total deviant verbalization, had an interjudge reliability of .96. Category means and standard deviations show only one instance of a major deviation between judgments — in the *E1* subcategory, which yielded a coefficient of .81, the lowest of all.

The data in Table II demonstrate that scores derived from the Sargent Test were not particularly stable from one administration to the next, yielding retest coefficients of reliability of from .44 to .76. However, they are probably as high as one could expect from a projective test and probably high enough to permit relationships to emerge if they exist to any significant degree. McClelland, Atkinson, Clark, & Lowell (1953, pp.193-194) have observed that a

"high product-moment correlation between measures on the same subjects on two occasions requires that the conditions of stimulation, whether they may be for each subject on the first occasion, be duplicated for him on the second occasion. . . . In addition to changed motivation states . . . the test-retest unreliability of the measure may be due to the change in the subjects produced by the first administration of the test. That is, it is theoretically possible to have a test which will correlate highly with a number of other measures (high "validity") but not with itself on a second administration (low "reliability"), if the first administration has somehow "spoiled" the subjects for this type of test . . . it should be considered as a possible handicap of projective tests."

Provided they are considered neither representative of enduring personal characteristics, nor suitable for precise prediction about the standing of individuals, the Sargent Test scores would appear sufficiently stable for group comparisons of the mode of interpretation of problem situations.

It is noted that both the interjudge and test-retest reliabilities for the 15 judged categories of the Sargent were considerably higher than those obtained by Sargent (1953) and Fassett (1948) for different categories. It is suggested that the higher reliabilities in the present investigation resulted from: (a) the requirement that all subjects respond to identical problem situations, which Sargent and Fassett did not include; (b) the considerably more explicit and detailed scoring procedure.

The rather close agreement between the scores assigned by judges of differing training and experience seems to support the contention that scoring of the Sargent Test may be accomplished with considerably greater objectivity than has been characteristic of projective tests. The Sargent Test, described by Anderson and Anderson as "one of the more promising of the newly proposed methods . . ." (1951, p.517), would seem to merit more attention as a diagnostic, clinical, and research tool. It is hoped that the scoring rationale herein presented will enable subsequent researchers to concentrate upon validation of the Sargent Test. It may be that this effort will encourage, correspondingly, further refinement of the scoring procedures.

REFERENCES

- Anderson, H. H., & Anderson, Gladys L. (Eds.) *An introduction to projective techniques*. New York: Prentice-Hall, 1951.
- Fassett, Katherine K. A preliminary investigation of the Sargent Test. *J. clin. Psychol.*, 1948, 4, 45-56.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
- Sargent, Helen D. *The insight test*. New York: Grune & Stratton, 1953.
- Schuman, E. P. Self-acceptance as a determinant of the mode of interpretation of problem situations. Unpublished doctoral dissertation, Columbia Univer., 1958.

Received April 9, 1962.

A Study of Some "Test Dimensions" Conceptions

WILLIAM SEEMAN
University of Cincinnati

PHILIP A. MARKS
Kansas University Medical Center

Among the "dimensions" or continua along which psychological tests can be ordered three which have been accorded considerable significance in the literature (especially the literature of projective techniques) have been selected as the subject of study in this paper. These are "stimulus ambiguity," "response freedom," and "depth."

The importance of stimulus ambiguity is too widely known to require any extensive comment. A statement like the following may be cited as representative: "... the subject is presented with a number of ambiguous stimuli and is then invited to respond to these stimuli. By such means it is assumed that the subject projects his own needs and press and that these will appear in responses to the ambiguous stimuli" (Bellak, 1950, p. 9). Although we do not intend to make much of it for the purposes of this paper we cannot refrain from calling attention to some considerable difficulties inherent in the very conception of "stimulus ambiguity." For some writers this appears to be synonymous with lack of stimulus "structure." (e.g., Stone and Dellis, 1960 p. 336). For others this appears not quite to be the case (e.g., Gibson, 1960 p. 699). At any rate, Gibson reports that in his survey of the literature he has been unable to turn up any definition for "structure" of a stimulus, while at the same time commenting that some experimental psychologists "would not agree that an inkblot is in any sense an unstructured stimulus" (1960, p. 699).

Response freedom is defined in terms of what the rules of the test permit. Thus, Hutt (1945) states

that in projective tests of personality (which are also called unstructured tests) "There is no restriction of the subject's use of the stimulus or his response to it." This is in contrast to the structured test in which the subject has "no opportunity of organizing in his own unique manner his response to the questions" (p. 136). It should be noted that while such rules provide *opportunity* for wide dispersal of responses they do not guarantee such response dispersal. For example, if one were to hand a group of subjects a full page advertisement of an automobile and ask "What could this be?" one might well expect a very narrow range of responses. Such response restriction, however, does not arise from the rules governing what the subjects are permitted to do; rather, it is the result of what is usually described as the "unambiguous" character of the stimulus. After all, in how many ways *can* one specify an automobile?

The "depth" dimension is related to the notion of "availability to consciousness." Thus, Stone and Dellis distinguish tests that "measure *primarily* 'surface' or 'depth' material" (p. 338). While the concept may also bear some relation to Cattell's distinction between "surface" and "source" traits, this does not appear to be exactly the same distinction. The concept of "depth" would also appear to be involved in a comment such as Murray's to the effect that special value resides in the power of the TAT to "expose the underlying inhibited tendencies" which the subject "cannot admit because he is unconscious of them" (1943, p. 1). We shall dwell no more on the extraordi-

FIGURE 1.

STIMULUS AMBIGUITY	RESPONSE FREEDOM	"DEPTH"	VERBAL LOADING
— Szondi	— Rorschach	— Word-association	— Word-association
— Strong	— Bender-Gestalt	— Sentence- completion	— TAT
— TAT	— Szondi	— Strong	— Bender-Gestalt
— Rorschach	— Sentence- completion	— Bender-Gestalt	— Strong
— Bender-Gestalt	— Word-association	— Rorschach	— Szondi
— Sentence- completion	— Strong	— TAT	— Rorschach
— Word-association	— TAT	— Szondi	— Sentence- completion

nary difficulties posed by attempts to specify the defining characteristics of the "depth" concept.

We turn now to a consideration of the questions this study is designed to answer. How are specified well known and widely used tests conceived to rank on these "dimensions?" What are the dimension intercorrelations? What tests are conceived to differ and what tests are not conceived to differ on these dimensions? Incidentally, our data seem to have some relevance for the hypotheses in which Stone and Dellis were interested: "Our specific hypothesis in regard to psychological tests," they state, "is that there is an inverse relationship between the degree of stimulus structure inherent in the test and the level of personality or impulse-control system being 'tapped' by the test. In other words, the more highly structured the test, the more likely that the great amount of data gained will be from a more conscious level of the personality and that the less structured the test, the less conscious the material obtained" (p. 336). In our terminology this asserts a significant correlation between "stimulus ambiguity" and "depth." Subsequently, Stone and Dellis refer to the inverse correlation between "stimulus and response structure" and the level of personality tapped (p. 339). Their conception of "response structure" appears to be identical with our conception of response freedom; i.e., less response structure is identical with greater response freedom.

In addition to answering these questions we were interested in one more aspect of the psychological tests

we elected to study: the conception of "verbal loading." Presumably, a test may be said to have a "high" verbal loading if its execution depends heavily on words (e.g., the Strong Vocational Interest Inventory) and a "low" verbal loading if it depends little or not at all on words (e.g., the Bender Gestalt).

PROCEDURE

The subjects were 36 highly experienced Ph.D. clinical psychologists from various parts of the country.¹ They were presented with the tests arranged as in Figure 1. The instructions required the subjects to rank order the tests on each "dimension," assigning the rank 1 as the "highest" rating for the specified dimension. Thus, the test judged to contain the "most ambiguous" stimuli was to be assigned a rank of 1 on the "SA" dimension; the test judged to allow "most response freedom" was to be assigned a rank of 1 on that dimension, the test judged to tap the "deepest" personality levels was to be assigned rank 1 on the "depth" dimension, and so forth.

RESULTS

It would, of course, have been surprising if the ranks for even one of the tests had been distributed in a "chance" fashion in the sense that each of the seven ranks had been assigned to any test with the same frequency. No such surprise occurred, all chi-square values testing such an equal distribution of frequencies be-

¹ Although responses were solicited and obtained also from non-Ph.D. clinical psychologists, these were not included in the analysis.

ing significant at the .01 level or beyond. That is, if each of the rank values one through seven is considered a cell, and the actual distribution of frequencies in the seven cells is tested against a theoretical equality for all seven cells, then all chi-squares are significant well beyond the .01 level.

For each of the tests the mean rank was then computed for the respective dimensions. The tests were then rank-ordered with respect to these mean ranks, and Spearman rank-order correlations were computed between dimensions. These are presented in Table I. Two of the correlations

TABLE I—Spearman Rank-Order Correlations Between the Specified "Dimensions"

	Response Freedom	"Depth"	Verbal Loading
Stimulus Ambiguity	.91**	.86*	NS
Response Freedom		1.00**	NS
"Depth"			NS

**Significant at the .01 level

* Significant at the .05 level

were significant at the .01 level and one at the .05 level. We confess to being more than moderately surprised at a perfect correlation between "depth" and response freedom. We can only plead that we are calling the cards as they actually fell! The hypothesis advanced by Stone and Dellis concerning the inverse relationship between "the degree of stimulus structure inherent in the test" and the "level of personality" which the test is supposed to tap is apparently a hypothesis held by these clinicians as well; and this is reflected in the .86 correlation between "depth" and stimulus ambiguity. But there seems to be even a stronger conviction concerning the relationship between response freedom and the "depth dimension."

Tables II through V present Kolmogorov values and corresponding probability levels for pairwise comparison of all the tests on the specified dimensions.

These data suggest that the Rorschach, in general, is viewed as a test that stands alone, at any rate with respect

TABLE II—Kolmogorov Values for Pairwise Comparison of Seven Tests on Stimulus Ambiguity

	TAT	Szondi	Sent C	Bender	Word A	Strong
Rorschach	18**	18**	18**	18**	18**	18**
TAT		10**	10**	10**	14**	18**
Szondi			6	7	6	17**
Sent Compl				6	4	14**
Bender					4	12**
Word Assoc						16**

**Significant at the .01 level

* Significant at the .05 level

TABLE III—Kolmogorov Values for Pairwise Comparison of Seven Tests on Response Freedom

	TAT	Szondi	Sent C	Bender	Word A	Strong
Rorschach	14**	16**	13**	18**	11**	18**
TAT		15**	5	18**	3	18**
Szondi			14**	4	12**	11**
Sent Compl				13**	2	18**
Bender					15**	10**
Word Assoc						16**

**Significant at the .01 level

* Significant at the .05 level

TABLE IV--Kolmogorov Values for Pairwise Comparison of Seven Tests on "Depth"

	TAT	Szondi	Sent C	Bender	Word A	Strong
Rorschach	10**	11**	14**	16**	13**	17**
TAT		13**	10**	15**	10**	16**
Szondi			6	9	8	9*
Sent Compl				11**	1	11**
Bender					14**	6
Word Assoc						11**

**Significant at the .01 level

* Significant at the .05 level

TABLE V--Kolmogorov Values for Pairwise Comparison of Seven Tests on Verbal Loading

	TAT	Szondi	Sent C	Bender	Word A	Strong
Rorschach	14**	13**	16**	15**	8	9*
TAT		15**	6	17**	7	9
Szondi			16**	12**	14**	14**
Sent Compl				17**	9*	9*
Bender					17**	15**
Word Assoc						9*

**Significant at the .01 level

* Significant at the .05 level

to conceptions of stimulus ambiguity,² response freedom, and "depth." All pairwise comparisons with the Rorschach on these "dimensions" are significant at the one percent level. This is to say that beyond question the Rorschach stimulus material ranks overwhelmingly as the most ambiguous of the test materials; that in response freedom it ranks even above the TAT (and, of course, above the other tests);

² A comment made by one of the subjects of this study is worth noting because it bears directly on the difficult problem of defining stimulus ambiguity. The question was raised whether "ambiguity" should be rated in terms of the "objective" nature of the stimulus independent of the task or in relation to the task. The significance of this, it seems to us, lies in the fact that stimulus ambiguity, at least as it usually conceived of in connection with psychological tests, is task-relevant. It must be defined in response terms. For example, if one decided to use the Rorschach cards with the instructions, "Sort these into two piles, placing in one pile the cards you like and in the other the cards you dislike," then it is difficult to see how the notion of ambiguity would have any meaning. It is only when a question like "What could this be?" elicits diversity of responses that we have some notion of stimulus ambiguity. That this is so is suggested by the question: What would be the implication for the stimulus ambiguity of card 5 of the Rorschach if *all* subjects saw it as a bat?

and that it is rated in "depth" significantly ahead of the other instruments. The Rorschach is assigned a significantly greater verbal loading than the Szondi and the Bender and a significantly lesser verbal loading than the TAT and the Strong; the difference between the Rorschach and the Word Association test does not attain statistical significance. The Strong also stands alone on stimulus ambiguity and on response freedom at the other end of the scale; i.e., it is least ambiguous and most restricted with respect to response permitted. In general, it rates low in "depth", though it is not significantly different in this respect from the Bender. The Szondi, the Bender, the Sentence Completions, and the Word Association tests are *not* rated appreciably different in stimulus ambiguity. With respect to "depth" the Strong no longer stands alone, but is rated much like the Bender and the Szondi. The Sentence Completions and the Word Association tests are much alike in their assigned ratings of "depth". They tend to be in the middle range of the rank values. As rated by their clinical judges, Stone and Dellis found a difference between the Rorschach and the TAT with respect

to "accessibility of pathology," and we find a difference between these two tests as rated on the "depth" dimension.

SUMMARY

Seven widely used psychological tests were ranked by 36 highly experienced clinical psychologists (all Ph. D.'s) with respect to stimulus ambiguity, response freedom, "depth" of personality level tapped, and verbal loading. Stimulus ambiguity, response freedom, and "depth" ratings were found to be significantly correlated. Although the Stone and Dellis hypothesized inverse relationship between degree of stimulus structure and "level of personality" supposed to be tapped by the test is supported by a significant correlation between "depth" and stimulus ambiguity rankings, an even stronger relationship turns up between the "depth" and response freedom ratings, in the form of a perfect correlation between these ranks. On the other hand, verbal loading was found to bear no significant relationship to stimulus ambiguity, to response freedom, or to "depth." By the Kolmogorov test the Rorschach stood alone at the "high" end of the

stimulus ambiguity ranking, the response freedom ranking, and the rankings of "depth." By the same test of significance the Strong stood alone at the "low" end. The Rorschach is rated as "more verbal" than the Szondi and the Bender and "less verbal" than the TAT and the Strong. The Szondi, the Bender, the Sentence Completions, and the Word Association tests are *not* rated in significantly different ways with respect to stimulus ambiguity. A number of difficulties associated with the conception of stimulus ambiguity and of "depth" were discussed.

REFERENCES

- Bellak, L. On the problem of the concept of projection. In L. E. Abt and L. Bellak (Eds.) *Projective Psychology*. New York: Knopf, 1950.
- Gibson, J. J. The concept of the stimulus in psychology. *Amer. Psychol.*, 1960, 694-703.
- Hutt, M. L. The use of projective methods of personality measurement in army medical installations. *J. clin. Psychol.*, 1945, 135-140.
- Murray, H. A. *Thematic Apperception Test*. Cambridge: Harvard University Press., 1943.
- Stone, H. K. & Dellis, N. P. An exploratory investigation into the levels hypothesis. *J. prof. Tech.*, 1960, 333-340.

Received April 5, 1962

An Analysis of Two Components Entering Into Rorschach Reliability Values

HARRY STEIN¹

New Jersey Diagnostic Center, Menlo Park

Studies of the reliability of Rorschach scores have been criticized because of failure to take account of the contaminating factor of differences in number of responses (Cronbach, 1949; Eichler, 1951; Jensen, 1959). This criticism applies not only to absolute scores of W, M, A, etc. but also to percentage scores, $W\%$, $M\%$, $A\%$, etc. In the latter case, the division by R reduces but does not remove the residual correlation with R. Several methods have been suggested for dealing with this problem. Eichler recommended that an analysis of covariance be employed and Cronbach suggested that either a fixed number of responses be obtained from each S or that reliabilities be separately determined for records of different lengths, i.e. 15-20R, 20-25R, etc. The latter method would so reduce the effects of variable R from different subjects that its influence on reliability would become negligible. To date, the author has been unable to find any Rorschach reliability studies employing the methods recommended by Eichler and Cronbach. Holtzman *et al* have abandoned the Rorschach and developed 2 parallel sets of ink-blots, each containing 45 cards and administered so as to obtain 1 response per card (1961). One of the major reasons advanced for their abandonment of the Rorschach set was the variable R and its contamination of reliability values.

In an elaboration of the criticism raised by Cronbach, Jensen discussed the error involved in percentage scores (called index scores) and said:

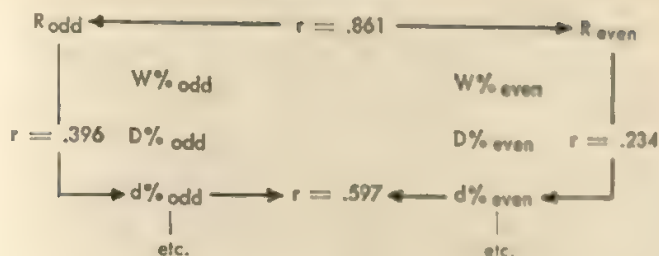
When an index score is a percentage of the total number of responses (e.g. $W\%$), its reliability will usually be spuriously high. The reason is twofold: there is usually a much greater range and variability in the denominator of the ratio than in its numerator (as in the case of W/R , where R is number of responses), and the reliability (at least the scoring reliability) is usually greater for the denominator, i.e. the number of responses. Most² of the index score reliability, therefore, would not be due to the reliability of the crucial element in the index, its numerator (e.g. W), but to the greater variability and reliability of its denominator (e.g. R), thus if R had a large standard deviation and was highly reliable, all other Rorschach scores could be made to appear *far more reliable*² than they actually are simply by estimating the reliability of their percentages: W/R , M/R , C/R , K/R , etc. This is a fallacious procedure (1959).

The above paper not only described the source of error involved in Rorschach index scores but concluded that most of the value of the reliability coefficient of correlation was the result of the spurious influence of R. A graphic illustration of the source of error introduced by the high intercorrelation of R with the split-half method of determining index score reliability is presented below. The intercorrelation of split half r 's is .861 which is then transmitted through the lateral correlations (R_{odd} has $r = .396$ with $d\%_{\text{odd}}$ and R_{even} has $r = .234$ with $d\%_{\text{even}}$). These side correlations then enter into the reliability correlation of $d\%_{\text{odd}}$ with $d\%_{\text{even}}$ yielding an $r = .597$. The above reliability values in the illustration as well as all data employed in this study were obtained from a previous reliability study of 53 Ss with physical disabilities (Stein, 1960).

¹I would like to express my thanks to Jane Mulhern, Josephine Spameni, Barbara Noll and Kathleen Quinlan for assisting with many of the calculations entering into the correlations.

²my italics

ILLUSTRATION



These previously obtained reliability values are repeated under Method A in Table I below.

To evaluate the amount of error contributed by the intercorrelation of R , it is necessary to reduce it to zero which then disrupts the chain of transmission. This was done with random tables in the following manner. Taking the total number of R given by the subject one entered into a column of random numbers and took the first number equal to or below the total R (Fisher and Yates, 1957). This number was then assigned to one side and the remainder to the other side. This division was then correlated to test for randomness and a value of $-.20$ was obtained. Although this value was not significant, such a relatively high value would prevent an accurate assessment of the error. Rather than repeat the random assignment for each side for an indefinite number of times until a value close to zero was obtained, corrections were made in two of the random assignments which then yielded an r of $.005$ between R for one side and R on the other side. It was then necessary to divide the responses of each record to yield the appropriate numbers of R on each side. This was done with random tables in a manner similar to the one described above. Taking the total R for the record the random tables were entered and all numbers equal to or below this were picked up until they added up to the necessary number of responses on one side of the random split. The remain-

der of the responses were placed on the other side. The reliability correlations obtained by this method after conversion of absolute scores into percentages are listed under Method C in Table I with an average r of $.226$.

TABLE I—Reliabilities of Rorschach Scores when Determined by Three Different Methods

Score	Method A	Method B	Method C
R	.861	.861	.005
W%	.571	.636	.421
D%	.256	.184	.097
d%	.597	.423	.398
Dd%	.266	.430	.351
S%	.205	.192	.407
M%	.139	-.014	.119
FM%	.117	.071	.169
m%	.358	.012	.330
F%	.359	.357	.282
k%	-.086	.203	.166
Fc%	.307	.065	.190
cF, c%	.074	.584	.407
FK%	.352	.052	.112
K, KF%	.057	-.056	.102
C'%	.308	.405	.331
FC%	.349	.219	.262
C, CF%	.323	.231	.180
H%	.087	.147	.152
A%	.516	.458	.260
Aobj%	.101	.122	.355
Anat%	.287	.418	.073
Obj%	.485	.470	.184
Nat%	.454	.520	.280
Plant%	.268	.050	.085
Geo%	.357	.584	.208
Art%	.463	.412	.541
Arch%	.562	-.065	.254
Cloud%	.261	.253	.177
Blood%	.000	-.019	.000
Fire%	.265	-.126	-.071
Emb%	-.069	.150	-.061
Average ^a	.288	.252	.226

^aThe averages were determined through 2 transformations

Table I, Method A lists the split half reliability correlations obtained in the usual manner of comparing the responses to the odd cards with the responses to the even cards. These yield an average reliability coefficient of .288. The difference between this value and the average coefficient of reliability of .226 through the method of randomizing R and responses (Method C) is the result of two components. The first component is the contaminating effect of the intercorrelation of R which has been removed by randomizing R. The other component is the influence of stimulus equivalence between the odd and even cards which has been removed by randomizing responses.

The next step in estimating the effects of these components was to keep the original number from R_{odd} and R_{even} ($r = .861$) as in the usual manner of determining reliability values (Table I, Method A) but to use the random tables to select the specific responses going under these numbers. This procedure eliminates the effects of equivalence between odd and even cards since we are comparing a randomly selected group of responses with the remainder of the responses. It should be remembered, however, that, since the correlation of R is still retained, the results reflect the contaminating influence of R. (One should also remember that chance error is present and cannot be evaluated. Since we have restricted our analysis to average reliability values, such error is not likely to be meaningful). The reliability results with removal of equivalence effects are presented under Method B, Table I. The average reliability coefficient was .252.

To recapitulate, the average reliability coefficient when employing the usual split half method is .288 (Method A, Table I). When the number of R_{odd} and R_{even} is held constant but the responses are randomly selected, the average reliability coefficient drops to .252. The reduction

from .288 to .252 results from the removal of the presumed stimulus equivalence between odd and even cards. When the total number of R assigned to each side of the split is randomly determined as well as a random assignment of responses, the average reliability is then further reduced to .226. The difference between .252 and .226 is the effect of contamination of reliability values by the "residual r with R".

DISCUSSION

The small reductions in average reliability value when one removes the influence of presumed stimulus equivalence between odd and even cards and then further removes the contaminating effects of "residual r with R" indicates several things. The difference between average reliability coefficient of .288 for odd-even split and .252 for average reliability when a random number of Rorschach responses are compared with the remainder of responses (Method B, Table I, where the numbers of R_{odd} and R_{even} are held constant) indicates that the balancing of stimulus values by comparing odd versus even cards adds little to reliability values. On the one hand, one might reason that the comparison of odd-even cards is a compromise between all stimulus values inherent in the cards and that the odd-even split does not fully reveal this stimulus equivalence. A true test of stimulus equivalence would then be an empirical evaluation of which split produces greatest consistent reliability values for each Rorschach score and that the splitting of cards could be different for each score. On the other hand, one might reason that responses to the Rorschach cards are determined much more by internal factors than by the stimulus values of the cards. In this case, of course, reliability values will not be materially improved no matter which way the cards are split. Evidence in support of this latter alternative is somewhat stronger at this point. Baugh-

man's (1958) study of responses to Rorschach cards with altered stimulus characteristics and his review of the literature indicate that certain kinds of alterations of the cards do not materially affect the scores. Stein's comparison between reliability values when measured by the split-half method and alternate response method disclosed a small decline in average reliability value with the alternate response method. This also suggests, that the largest part of the individual's consistent pattern of responding is determined by inner determinants rather than by the stimulus values of the blots. These inner determinants would include the set generated by instructions and the first card presented to the S. However, these alternate points of view are not conclusively supported from studies presently available.

The reduction from average reliability of .252 (Method B, Table I) to .226 when both R and responses are randomized to remove the contaminating influence of R (Method C, Table I) indicates that the inflation of reliability values by the traditional method of determining Rorschach reliability is rather small. There is no doubt that this inflation does occur but it is not so serious as Jensen (1959) has stated. Instead of "most" of the reliability coefficient reflecting the contaminating influence of R, only a very small part of its value can be attributed to this factor.

CONCLUSION

An evaluation and analysis of Ror-

schach reliability correlations of 31 scores (W%, D%, etc.) discloses that removal of the effects of stimulus equivalence between odd and even cards lowers the average reliability from .288 to .252. Removal of the contaminating effects of the "residual r with R" upon the reliability correlations by means of randomizing R on each side of the split and randomizing the assignment of response further lowers the average reliability coefficient from .252 to .226. This small decline indicates that the inflation of reliability values by the "residual r with R" is small and does not appreciably lower the values previously reported in the literature.

REFERENCES

- Baughman, E. E. The role of the stimulus in Rorschach responses. *Psychol. Bull.*, 1958, 55, 121-147.
- Cronbach, L. J. Statistical methods applied to Rorschach scores: A review. *Psychol. Bull.*, 1949, 46, 393-429.
- Eichler, R. M. Some comments on the controlling of differences in responses on the Rorschach test. *Psychol. Bull.*, 1951, 48, 257-259.
- Fisher, R. A. & Yates, F. *Statistical tables for biological, agricultural and medical research*. New York: Hafner Publishing Co. Inc., 1957.
- Holtzman, W. H., Thorpe, S. T., Swartz, J. D. and Herron, E. W. *Inkblot perception and personality*. Austin: University of Texas Press, 1961.
- Jensen, A. R. The reliability of projective techniques: Review of the literature. *Acta Psychologica*, 1959, 16, 3-67.
- Stein, H. Rotation and reliability of the Rorschach. *J. proj. Tech.*, 1960, 24, 171-181.
- Stein, H. An evaluation of Rorschach reliability through the alternate-response method. *J. clin. Psychol.*, 1961, 17, 241-245.

Received March 30, 1962

Examiner Influence on Thematic Apperception Test Responses

GEORGE C. TURNER AND JAMES C. COLEMAN
University of California, Los Angeles

A number of investigators have pointed to the importance of examiner differences in determining responses to projective tests (Alden & Benton, 1951; Baughman, 1961; Jensen, 1959; Lord, 1950; Murstein, 1959; Sarason, 1958; Summerwell, 1958). However, most of the studies in this field have been concerned with the Rorschach test. The present investigators were able to find only one study that dealt with the influence of the examiner on Thematic Apperception Test responses, and this study was primarily concerned with differences in test instructions (Summerwell, 1958). Although only two examiners were involved, a difference due to examiner influence in terms of outcome of stories was found for one of four sets of instructions—indicating that examiner differences may be important in influencing TAT responses.

Since the TAT is a widely used projective instrument for both clinical diagnosis and research use, a more definitive evaluation of examiner influence on TAT protocols would appear to be indicated. The present study represented an attempt to help fill this research gap. Two specific objectives were involved: 1) to ascertain whether examiners influence the amount or type of TAT responses elicited from the subjects they test, and 2) to delineate examiner characteristics related to such response differences if they do exist.

METHOD

The key elements of the present experimental design may be summarized as follows.

Subject and examiners.

Two different populations were employed in this study, the subject

population and the examiner population. The subject population whose TAT records were used in this investigation consisted of 204 male outpatients who were receiving treatment at a Veterans Administration Mental Hygiene Clinic. These patients ranged in age from 20 to 47 years with a mean of 30 years. A wide range of diagnostic categories were represented, but the majority of patients were classified as neurotics. Although all of these patients were able to function outside of a mental hospital, the degree of maladjustment ranged from mild to severe in terms of their ability to function socially and vocationally.

The examiners who administered the TAT consisted of 24 third and fourth year graduate students in clinical psychology. These students came from two universities and were assigned to the Veterans Administration for their internship training. Twenty of the examiners were male and four were female. Training requirements insured that each examiner had reached a fairly high degree of skill in the use of projective techniques including the TAT.

Since the assignment of patients to examiners was done on a rotational basis, it was assumed that the patient sample was randomized as far as the examiners were concerned. It is also relevant to note that this study utilized test material which had been administered in a clinical setting without the examiners' awareness that the test results would be used for research purposes. It was felt that this method of collecting data had the advantage of minimizing the possible effects of the set an examiner might assume if he knows that he is administering a test under experimental conditions.

TAT variables.

The problem of measuring the amount and type of productivity in TAT protocols was resolved by the use of two types of variables—thought units and affective units.¹ These variables appeared appropriate to the type of material obtained on the TAT and had the further advantage of minimizing the fact that few examiners write down the subject's stories word for word, but usually do manage to record key ideas and feelings.

Since thought units may vary in type as well as in number, three categories of thought units were distinguished: 1) Picture description TU's (P)—thought units which enumerate or name things in the picture without connecting them to a story, 2) Storytelling TU's (S)²—thought units which interpret the picture, e.g., portray action, express thoughts and feelings of characters, or deal with events pertaining to a story made up about the picture, and 3) Irrelevant TU's (I)—thought units which are not relevant to the picture or a story made up about the picture, e.g., expressions of inadequacy or self reference such as "I can't tell stories."

Two scoring categories were utilized for affective units—Warmth (W) and Hostility (H). These two variables were scored independently of the P, S, and I variables, but on the basis of the thought units which expressed them. Thus a thought unit could be scored P, S, or I with or without an additional W or H score. Hostility (H) was scored only if hate, resentment, dislike, rejection, criticism, and related feelings were expressed openly in word or behavior. Warmth (W) was scored if love, affection, acceptance, reassurance, praise, or related manifestations were expressed openly in word or behavior. Both Hostility and Warmth could be directed toward an object, concept,

institution, or person. Neither Hostility nor Warmth were scored if they were covert or required an interpretation by the examiner.

In scoring a TAT protocol, the meaning of a thought unit was considered in relation to the total context of the description or story and not in isolation. Sometimes it was necessary to relate a given TU to preceding and succeeding TU's to obtain the true meaning and score it properly.

In order to insure the reliability of the scoring system, a sample of 72 stories were evaluated by two judges working independently. The results showed a very high degree of agreement for both the total number of TU's ($r = 0.96$) and for the five scoring categories ($T = 0.89$).³

Examiner variables.

The isolation of examiner variables which might influence the subject's TAT responses was dictated by both research findings relating to other projective tests and by practical considerations. The variables which were finally selected were: 1) Experience in the use of the TAT, 2) Preference for the TAT, 3) Personality traits of Dominance, Warmth, Hostility, and Shyness, and 4) Diagnostic Competence.

Data concerning these examiner variables were collected in a number of ways. The examiner's experience in using the TAT was ascertained by having each examiner fill out an Information Form indicating the number of years of experience he had had in administering the TAT, and by tabulating the total number of TAT's each examiner had administered in the VA Mental Hygiene Clinic. Preference for the TAT was ascertained by having each examiner indicate a preference rating for the TAT in relation to five other commonly used diagnostic tests and by the ratio of

¹The concept of thought-unit or TU was developed by Dollard and Mowrer (1947) and is used here with certain modifications.

²Tschuprow's coefficient (T) as given by Yule and Kendall (1947, p. 54-55) was employed. This coefficient is essentially a correction of the contingency coefficient when using a small number of categories (1947).

diagnostic cases in which the TAT was used to the total number of cases tested by the examiner in the VA Mental Hygiene Clinic. The personality traits of Dominance, Warmth, Hostility, and Shyness were measured by rating scales in which each examiner rated the other examiners that he felt he knew well enough to assign valid ratings. In addition, each examiner was encouraged to rate himself on these traits. Diagnostic Competence was also measured by rating scales following the same procedure indicated for measuring personality traits. (The rating scale for Diagnostic Competence was based on the definition suggested by Kelly and Fiske (1951).

Since most of the examiners knew each other and had worked together over a period of several years, it was assumed that they could rate each other on the dimensions indicated with acceptable reliability and validity. Intraclass correlations (r') utilizing a sample of 15 raters (a sample of the ratings made by 15 raters on thirty examiners on all of the scales), yielded correlations (r 's) ranging from 0.78 to 0.88 indicating a high level of agreement among raters re use of (r') (Garrett & Zubin, 1943).

Procedure.

Much of the information concerning the procedure was indicated in the discussion of TAT and examiner variables. However, a few additional details will help to complete the picture.

During the entire course of data collecting a close watch was maintained in order to screen out TAT records or examiners where artifacts might operate to influence the testing situation or results. Thus all protocols were eliminated where it was known that the patient had taken the TAT before; similarly protocols were eliminated in cases where the testing procedure was altered in some fashion, such as the administration of

the test under experimental conditions.

In the process of data collecting one key problem arose when it was discovered that not all of the TAT cards were administered to each patient. Typically each examiner either administered a favorite selection of cards or administered the cards which he felt were most pertinent to the case in question. Only cards 1 and 13MF were administered with sufficient regularity to retain most of the examiners and a majority of the subject population. As a consequence the final selection of stories utilized only cards 1 and 13MF. Although this selection was made out of necessity, it turned out to be a fortunate choice. Card 1, as the introductory picture, obtains the subject's initial reaction to the test and perhaps reflects the examiner-patient relationship more than some of the other cards. Card 13MF, with its strong emotional and sexual elements would in turn appear to put the subject-examiner relationship to the acid test.

The final selection of TAT data consisted of 408 stories told by 204 subjects to 24 examiners. The number of subjects per examiner ranged from a minimum of five to thirteen, with a mean of 8.5 subjects per examiner. For each story the number and type of thought units were recorded in accordance with the method described earlier in this paper.

In the construction of rating scales for assessing Diagnostic Competence and personality traits of examiners, four discrete categories to which numerical values could be assigned were utilized instead of a continuous distribution. This was intended to simplify the rating task of making each category as explicit as possible.

As indicated previously each examiner was asked to rate only those other examiners that he felt he knew well enough to assign valid ratings. This yielded sixteen or more ratings for each of the examiners. To avoid influencing the raters unnecessarily,

they were given only a very general idea of the purpose of the study, but the need for their cooperation if the experiment was to prove of value was emphasized. The examiners were also asked to make their ratings in the order presented and not to refer back or to look ahead — a procedure designed to minimize the possibility of ratings on one scale influencing their ratings on another.

Statistical methods utilized in the treatment of the data will be indicated in the next section dealing with the results of the present investigation. However, it may be pointed out that non-parametric statistical procedures were used, since the data did not satisfy the requirements for homogeneity of variance essential for the use of parametric statistical procedures (Edwards, 1950).

RESULTS AND DISCUSSION

The results of the present investigation focus around two key questions: 1) do examiners differ in the amount or type of TAT material they elicit from subjects, 2) if differences do occur in the amount or type of material elicited, are these differences related to the examiner variables under investigation?

Examiners and TAT material elicited.

In determining whether differences exist among examiners in the total amount or type of material elicited from subjects, a "distribution-free" method, suggested by Mood (1950) was employed. This method is some-

what similar to a simple analysis of variance since it tests the probability that two or more samples are drawn from the same population or from populations having the same median.

The results indicated that the examiners did not differ at an acceptable level of significance with respect to the total number of TU's elicited. However, the examiners were found to differ significantly in the number of Story-telling (S) thought units elicited ($P = .05$). A tendency toward significance of difference among examiners was reflected in the "P" and "I" production, with the $P = .06$ in each case. With respect to AU's a significant difference was found in the case of Warmth Units ($P = .01$) but not with respect to Hostility Units. It may be recalled that AU's are TU's given an affective scoring—H or W—after being scored as a TU. Thus in all but the production of Hostility Units, a qualitative breakdown brought out significant or near significant differences among examiners.

Examiner variables and differences in TAT responses. The influence of examiner variables on TAT production was assessed by several different procedures. 1) Experience. The examiner's experience with the TAT—in terms of years of experience was obtained directly from the information form. The examiners were then divided into three nearly equal groups: those with two to three years experience, those with three and one-half to four years, and those with four and one-half or more. The re-

TABLE I—Differences Between Examiners in the Amount of Production on the TAT, Measured in Terms of Thought Units (TU's).

Thought Units	N	Result	Degrees of Freedom	P
Total number of TU's.	24	$X^2=31.48^a$	23	0.11
Story-telling TU's.	24	$X^2=36.91$	23	0.04
Picture description TU's.	24	$X^2=34.51$	23	0.06
Irrelevant TU's.	24	$X^2=34.12$	23	0.06
Hostility (h) AU's.	24	$X^2=29.64$	23	0.16
Warmth (w) AU's.	24	$X^2=40.64$	23	0.01

^a Significance of difference is tested by Mood's (1950) Distribution-free Test, which is equivalent to chi square.

sults — utilizing Mood's method — ranged from $P = 0.14$ to 0.62 indicating that the years of experience in giving the TAT were not related to the type of material elicited from subjects. A second measure of the examiner's experience with the TAT was a tally of the number of TAT's administered during his stay at the VA Mental Hygiene Clinic. The examiners were then divided into three groups; those who had administered six or seven TAT's, those who had administered eight or nine, and those who had administered ten or more. Mood's method was again used. The results show that both the "S" and "P" TU's show differences significant near the 0.01 level of confidence. The examiners who gave fewer TAT's obtained more "S" TU's while those who gave ten or more TAT's elicited more "P" TU's.

2) Preference. The examiner's preference for the TAT in relation to type of material elicited from subjects was assessed next. On the information form the examiners were asked to distribute 100 points among six tests, the MMPI, TAT, Sentence Completion, Shipley-Hartford, Rorschach, and Wechsler-Bellevue as a means of signifying their preference for the general use of each test. A ranking for each of the examiners, in terms of their preference among these tests was obtained, and those who ranked the TAT first or second were put in one group designated as the "high preference group." Those who

placed the TAT lower than second place were put in a second group designated the "low preference group." Mood's method of difference between medians was again applied. The probability levels that the medians do not differ significantly range from 0.025 to about 0.80. Only in one category, the number of Hostile Units elicited, did the level of confidence go beyond the 0.05 level. In this case the examiners showing "high preference" elicited more Hostile Units than those showing "low preference" for the TAT. A second approach to the examiner's preference for the TAT utilized the ratio of TAT's administered to the total number of cases tested at the VA Mental Hygiene Clinic. Such a ratio reflects the relative frequency that an examiner uses the TAT, holding the number of cases the examiners test constant. This ratio probably shows a combination of degree of preference, interest, and degree of usefulness the examiner finds in the test. The examiners were divided into thirds by placing those with ratios of less than 0.60 in one group, those with ratios of 0.60 to 0.79 in the second group, and those with ratios above 0.79 in the third group. The results—utilizing Mood's method—ranged from $P = 0.28$ to 0.58 . Thus the ratio of TAT's given to the number of cases tested by the examiner was not significantly related to the types of response elicited.

3) Personality traits. The next examiner variable assessed was the re-

TABLE II—Summary of Tests of Relationship Between Amount of Production^a on the TAT and Examiner Variables

Examiner Variable	Total TU	Story-Telling	Picture-Description	Irrelevant	Hostility	Warmth
Years experience with TAT	0.17	0.42	0.62	0.28	0.42	0.14
Number of TAT's administered	0.27	0.014*	0.011*	0.76	0.31	0.76
"High" or "Low" preference	0.58	0.58	0.22	0.80	0.025*	0.80
Number TAT's/Total Cases tested	0.58	0.28	0.36	0.42	0.58	0.28

^a The amount of production is measured in terms of the mean number of Thought Units (TU) per protocol for each type of TU designated. The numbers in the above table are probability levels. Mood's (1950) "distribution free" method was used, with an N of 23 examiners.

* Significant at or beyond the 0.05 level.

lationship of the personality traits of Dominance, Warmth, Hostility, and Shyness to the responses elicited on the TAT. The influence of each of these traits on TAT production was evaluated by first computing the mean rating given each examiner by all the judges who rated the examiner. The examiners were then ranked on each scale according to these mean ratings. The statistical method applied to test the significance of differences for these data is one suggested by Wilcoxon (1949) for use with unpaired replicates. Using this method the eight examiners rated highest and the eight rated lowest on a scale were ranked according to the number of units of a particular type they elicited. The sums of the ranks for each group were then obtained. Wilcoxon provides a table of smaller rank totals necessary for significance at the .05, .02, and .01 levels. For an N of eight, which is applicable here, the smaller rank total must be 49 or less to be significant at the .05 level. All of the rank totals exceeded 49 indicating that personality traits of examiners were not significantly related to types of TAT responses.

Although the differences did not reach statistical significance, it is interesting to note that those examiners who were rated highest in Warmth elicited more Hostile Units than those examiners who rated low in Warmth. Here it would appear conceivable that a warm, permissive examiner might create a testing atmosphere which would permit the subject to show the hostility he really feels, rather than forcing him to cover it up with an assumed or defensive warmth. Curiously enough, however, the "warm" examiners elicited fewer "S"'s than the examiners who rated low in Warmth, although again the differences were not statistically significant.

4) Diagnostic Competence. Diagnostic Competence as a possible factor influencing TAT production was the final examiner variable considered. The method of testing the significance of this variable was the same as that employed with the personality variables, namely, Wilcoxon's test for significance of differences between unpaired replicates. Again the eight examiners rated highest and the eight rated lowest in Diagnostic Competence were ranked according to the

TABLE III—Relationships Between the Amount of Production (thought units) and Examiner Variables (ratings)

Variable	Type of Thought Unit					
	Total TU's	Story Telling	Picture Description	Irrelevant	Hostility	Warmth
Dominance						
Eight Highest	66.5 ^a	72.0	62.5	66.5	77.0	70.0
Eight Lowest	69.5	64.0	73.5	69.5	59.0	66.0
Warmth						
Eight Highest	61.5	59.0	70.5	66.0	72.5	59.5
Eight Lowest	74.5	77.0	65.5	70.0	63.5	76.5
Hostility						
Eight Highest	64.5	71.0	59.0	59.5	67.0	68.5
Eight Lowest	71.5	65.0	77.0	76.5	69.0	67.5
Shyness						
Eight Highest	64.0	63.5	57.5	65.5	60.5	66.0
Eight Lowest	72.0	72.5	78.5	70.5	75.5	70.0
Diag. Comp.						
Eight Highest	75.5	77.0	81.0	75.0	71.5	70.5
Eight Lowest	60.5	59.0	55.0	61.0	64.5	65.5

^aWilcoxon's (1949) Significance of Difference between Unpaired Replicates

The smaller sum of ranks must equal 49 or less for significance at the 0.05 level. None of the above rank sums achieved this level of significance. See text above under Personality Traits for an explanation of the statistical treatment.

number of units of a particular type they elicited. The sums of the ranks for each group were then obtained. All of the smaller rank totals exceed 49, indicating that the Diagnostic Competence of the examiner—as assessed by ratings of his peers—does not seem to effect production on the TAT. However, those examiners whose mean ratings placed them among the eight highest in Diagnostic Competence ranked higher in the amount of production they elicited from the subjects tested than those examiners whose mean ratings placed them among the eight lowest in Diagnostic Competence. This situation held for both TU's and AU's of all types. However, we are speaking of a trend here and not statistically significant results.

Of particular importance to the present investigation is the finding that examiners do differ in the type of responses which they elicit from subjects on the TAT. Some examiners elicit significantly more Story-telling and Warmth Units than do others. Since Story-telling and Warmth Units are considered of more value in assessing personality traits than are Picture-description and Irrelevant thought units, such a difference among examiners would appear relevant in evaluating the reliability and validity of the TAT as an assessment instrument. If some examiners are able to elicit material of more diagnostic significance than others, this is likely to be reflected in the quality of the interpretations made by different examiners and in the total diagnostic description of the subject. Thus the present results support Jensen's conclusion that even if scoring, retest reliability, and internal consistency were all perfect for scores obtained by one examiner, the reliability of the test might be low when it is used by a number of examiners (Jensen, 1959). It would appear important to either control or eliminate examiner influence when the TAT is being used for clinical or research purposes.

The results relative to examiner variables that might be related to differences in productivity of subjects on the TAT are not so clear-cut nor conclusive. It was hoped at the beginning of this investigation that a general personality description of a "productive" examiner might be possible in terms of the examiner variables under investigation. However, positive results here were thinly sprinkled among negative results. Nevertheless it is of interest to note that examiners who had less experience with the TAT elicited significantly more Story-telling TU's than those who had more experience. It would appear possible here that as an examiner gains familiarity with the TAT, he may make less effort and take less care in his administration of the test. It is also interesting to note that examiners showing "high preference" for the TAT elicited significantly more Hostile Units than those showing "low preference" for this instrument—as indicated by their ratings of the TAT in relation to five other commonly used clinical tests. The meaning of this finding is not clear.

The failure to find a significant relationship between examiner traits of Dominance, Warmth, Hostility, Shyness and Diagnostic Competence indicates the need to redefine these variables or to introduce new ones for investigation. Since it appears probable that the rating scale method produced reliable and valid results, it would appear that examiner variables other than those considered in the present experiment should be assessed for their possible influence. The findings of Summerwell et al., would also indicate the importance of studying the instructions used by given examiners in administering the test—e.g., these investigators found that Murray's instructions resulted in more depressive, sadder stories than did neutral instructions (Summerwell, 1958). A means of evaluating combinations of examiner variables instead of single variables might lead

to more definitive results. This is based upon the assumption that in general people react to others as they perceive them *in toto*—as whole persons—and not to single characteristics, although of course a single characteristic may be very important. Moreover, it would appear relevant to consider the relationship between the subject and examiner as a dynamic and possibly changing one. Thus examiner traits that influence certain types of patients may be less important for other types, e.g., patients suffering from neurotic disorders may be influenced by certain examiner characteristics while those diagnosed as having character disorders may be influenced by other examiner characteristics. In short the characteristics of the patient as well as the examiner may have to be considered. In addition, it is possible that examiner variables that are important early in the testing situation may not be important during the latter part of the testing. In any event, the present study does point to the need for further research in this area.

SUMMARY

Two general questions were central to the present investigation: 1) do examiners influence the amount or type of TAT responses elicited from the subjects they test, and 2) what examiner characteristics are related to such response differences if they do exist?

The subject population whose TAT records were used in this investigation consisted of 204 male outpatients receiving treatment at a Veterans Administration Mental Hygiene Clinic. The examiner population consisted of 20 male and four female graduate students in clinical psychology who were assigned to the Veterans Administration for their internship training. Two types of variables were involved in the present investigation—TAT variables and Examiner variables. Thought units (TU's) and affective units (AU's)

were adopted as the basic TAT variables. Thought units were in turn divided into types: Story-telling (S) TU's, Picture-description (P) TU's, and Irrelevant-material (I) TU's. Affective units were scored independently of the S, P, and I variables and included Warmth (W) and Hostility (H). Examiner variables which were isolated for study in relation to amount and type of TAT productivity were: 1) Experience with TAT, 2) Attitude toward or preference for TAT, 3) Personality traits of Dominance, Warmth, Hostility, and Shyness, and 4) Diagnostic Competence. Experience with and attitudes toward the TAT was ascertained primarily by an Information Form given to each examiner. Personality traits and Diagnostic Competence were assessed by means of rating scales in which each examiner rated the others. The findings of the present investigation may be summarized as follows:

- 1) Examiners did not differ in the total amount of material—TU's—elicited from their subjects.
- 2) Examiners did differ in the type of material elicited—significant differences occurring in the number of Story-telling and Warmth units.
- 3) The number of years of experience in giving the TAT was not related to TAT responses.
- 4) Examiners showing a high preference for the TAT as a diagnostic instrument elicited more Hostile Units.
- 5) Examiners who had given fewer TAT's elicited more Story-telling material than those who had given more TAT's.
- 6) No statistically significant differences were found in the amount or type of TAT responses and the examiner traits of Dominance, Warmth, Hostility, or Shyness.
- 7) Examiner differences in Diagnostic Competence were not significantly related to the

amount or type of TAT productivity, although there was a tendency for those examiners rated highest in Diagnostic Competence to elicit more TAT material—both in terms of TU's and AU's.

The present results point to the importance of further investigating examiner variables in relation to subject's responses on the TAT both in relation to clinical usage and research usage of this projective instrument.

REFERENCES

- Alden, Priscilla & Benton, A. L. Relationship of sex of examiner to incidence of Rorschach responses with sexual content. *J. proj. Tech.*, 1951, 15, 231-234.
- Baughman, E. E. Rorschach scores as a function of examiner differences. *J. proj. Tech.*, 1951, 15, 243-249.
- Dollard, J. & Mowrer, C. H. A method of measuring tension in written documents. *J. adnorm. soc. Psychol.*, 1947, 42, 3-32.
- Edwards, A. L. *Experimental design in psychological research*. New York: Reinhart, 1950.
- Garrett, H. E. & Zubin, J. The analysis of variance in psychological research. *Psychol. Bull.*, 1943, 40, 233-267.
- Jensen, A. R. The reliability of projective techniques: review of the literature. *Acta Psychologica*, 1959, 16, 100-136.
- Kelly, E. L. & Fiske, D. W. *The prediction of performance in clinical psychology*. Ann Arbor: Univ. Michigan, 1951.
- Lord, Edith. Experimentally induced variations in Rorschach performance. *Psychol. Monogr.*, 1950, 64, No. 10.
- Mood, A. M. *Introduction to the theory of statistics*. New York: McGraw-Hill, 1950.
- Murstein, B. I. A conceptual model of projective techniques applied to stimulus variations with thematic techniques. *J. consult. Psychol.*, 1959, 23, 3-14.
- Sarason, Barbara R. & Sarason, I. G. The effect of type of administration and sex of subject on emotional tone and outcome of ratings of TAT stories. *J. proj. Tech.*, 1958, 22, 333-337.
- Summerwell, Harriet C., Campbell, Mary M. & Sarason, I. G. The effect of differential motivating instructions on the emotional tone and outcome of TAT stories. *J. consult. Psychol.*, 1958, 22, 385-388.
- Wilcoxon, F. *Some rapid approximate statistical procedures*. New York: American Cyanamide Co., 1949.
- Yule, G. U. & Kendall, M. G. *An introduction to the theory of statistics*. London: Griffin, 1947.

Received December 20, 1961

Revision received March 26, 1962

BOOK REVIEWS

Burgemeister, Bessie B. *Psychological Techniques in Neurological Diagnosis*. New York: Harper & Row, Publishers, 1962, \$7.50.

This 248 page volume, according to its author, is intended to bring together knowledge concerning both the values and limitations of psychological techniques for neurological diagnosis.

The general and most substantive part of the book is the first 60 pages. At the beginning there is a thorough and systematic review of current theory and knowledge concerning the psychological sequelae of neurological disorders. The difficulty in this area concerns the relatively picayune approach of neuropsychology with its concomitant difficulty in identifying large functional units as contrasted to the rather vague holistic approach of psychology with its reluctance to make specific predictions. The picture is confused by problems of definition and etiology. The task of neuropsychology is defined by the author as the detection of disturbances in the integrative capacity of the nervous system. The two main approaches open to psychologists are clinical observation and formal testing.

Observation is, in the author's opinion, used less than she considers desirable because of the psychologist's supposed "dread" of being considered unscientific. To aid the clinician in the increasing utilization of the observational technique, typical behavioral signs of organicity are given. However, such signs, just as test signs, suffer from vagueness, lack of direct behavioral referents, overlap among groups, and lack of specificity.

Burgemeister cites three classes of tests which she deems useful in the diagnosis of neurological disorders. The basic hypothesis in using intelligence tests is that organicity is measured by the estimated discrepancy between optimal and present functioning. The author cites much clinical folklore in this regard and adds some of her own hunches. A review of the literature in this area is rather limited and disappointing, especially since little of the rather overwhelming negative evidence concerning the utility of Wechsler-type scales for the detection of organicity is mentioned. The author goes on to list the various special tests that have been designed

for the detection of neurological disorders. These include the Shipley-Hartford Retreat Scale, the Hunt-Minnesota, the Halstead, the Bender-Gestalt, the Benton, the Graham-Kendall, and the Hooper Visual Organization Test. In addition to these, there are the Goldstein-Scheer techniques, including the Cube Test, Color-Sorting, Object-Sorting, Color-Form Sorting, and Stick tests. The review of the literature concerning these special tests is much more critical and comprehensive, with reference to much of the pertinent literature.

Burgemeister's reaction to the failure of quantitative approaches, as illustrated by the Shipley-Hartford, Hunt, and Halstead, is to espouse qualitative approaches such as the Hanfman-Kasanin, Zaslow and Goldstein techniques. Her motto seems to be that a devil unknown is better than a devil known. Whether it is really true that qualitative observations which cannot be proved invalid are superior to quantitative approaches of doubtful validity is of course, a matter of taste.

A discussion of projective methods is also a little disappointing. Under these she includes drawing techniques, the Mosaic, and of course, the Rorschach. In her discussion of the Rorschach she mentions the various lists of signs of which she seems quite enamoured. However, she does not mention some of the most frequent criticisms of the sign approach, such as the overlapping of groups, the lack of utility of signs based on extreme groups for borderline diagnoses, the frequent lack of behavioral referents, the unreliability of judgments, etc. In her view, the most valuable aspect of some of the better sign systems is that they have a certain amount of generality beyond their test source of information and can be conceptualized in gross behavioral terms as well.

In a brief section on the neuropsychological examination, Burgemeister makes some suggestions for conducting the examination and integrating the findings. Her suggestions seem sound, but are rather sparse, and could well be elaborated. Little emphasis is given to the necessity of diagnosing not only the nature and degree of the organic disorder, but also the patient's particular emotional reactions to it. Emotions in this book are thought of more as signs of organicity rather than reactions to it.

The discussion of various organic syndromes

that follows is of a pattern. In each instance there is a sophisticated discussion of the syndrome, its neurophysiological characteristics, and its psychological sequelae. This is followed by brief remarks concerning expected behavior on psychological tests which are in turn illustrated by case histories. Syndromes discussed include epilepsy, mental retardation (with specific reference to conditions with known organic etiology), cerebral palsy, traumatic head injury and intra-cranial neoplasms, psychosurgical problems, multiple sclerosis, and geriatrics. In each one of these sections, the author does her best. The general description of these syndromes and the suggestions as to test indicators should certainly be useful to any clinical psychologist or clinical psychology trainee working for the first time in a setting where such pathology is apt to occur. The little tangible evidence available for the predictive efficiency of test signs is cited, and no attempt is made to be excessively dogmatic on the basis of the few facts that exist.

The book appears to be somewhat padded with the many case illustrations and test protocols, but these will undoubtedly be of value to students and instructors in illustrating the various points made. The over-all effect of the volume is that of modesty rather than grandiosity. The little knowledge that we possess concerning the utilization of psychological tests and observations for the diagnosis of organic brain damage is nicely brought together between two covers and the book will be of great value in both training students and reminding experienced clinicians of these areas when they run across them only infrequently. No one should be disappointed if Burgemeister can only review work that has been done. Her admonition to emphasize qualitative observations in the detection of neuropsychological disorders strikes a responsive chord, but begs the question as to the validity of diagnosis based on such unquantifiable criteria.

WALTER G. KLOPPER
Univ. of Portland

Buros, Oscar K. *Tests in print*. Highland Park, N.J.: Gryphon Press, 1961. Pp. xxix + 479. \$7.00.

Buros' most recent volume is a handy index and supplement to his *Mental Measurements Yearbooks*.

By dint of persistent detective work with test publishers, the author has discovered and established which English tests are still pub-

lished and available for use and which tests have been published since his last yearbook (1959). These tests he has classified and documented with cross-references to his comprehensive yearbook summaries, bibliographies and reviews.

He is careful to specify, when it is possible to do so, the state of development of each test, i.e. whether it has been validated or is an experimental form. And he urges great caution in the selection and utilization of tests. Buros has devoted his professional life to the development, implementation and communication of psychological testing, information and methodology, and his volumes have added tremendously to the organization and structuring of our knowledge. Most importantly, he consistently reminds us of the limitations of tests and refuses to indulge our needs to select and use tests mechanically in order to avoid the time and effort of critical thinking and evaluation.

An added bonus is his inclusion of the *Technical Recommendations for Psychological Tests and Diagnostic Techniques* prepared originally by the APA Committee on Test Standards plus *Technical Recommendations for Achievement Tests* developed by the American Educational Research Association and the National Council on Measurements. Newcomers to clinical psychology are likely to be less familiar with these reports than they ought to be. Their republication gives this book a greater degree of integration and autonomy than it would have had without them.

The book has several unique features that make it virtually indispensable to persons in counselling and educational measurements as well as to the clinician who performs research or who wishes to increase his diagnostic armamentarium with both projective and non-projective techniques, in order to deal with special issues not adequately assessable with global psychodiagnostic techniques. It lists all known available tests, experimental as well as standardized. It provides references to reviews and bibliographies. It indicates existing variants of well-known tests. Hence it is essentially the only source to which one can go in search of specialized techniques. It includes an exhaustive directory of test publishers and distributors, foreign as well as domestic, a title index of the tests and an index of test authors.

What more could one wish for in a catalog of tests? And why in the world would anyone who uses tests not want a copy?

BERTRAM R. FORER
Los Angeles, Calif.

Harrower, Molly. *The practice of clinical psychology*, Springfield, Ill.; C. C. Thomas, 1962.

This book is winning, fascinating, and ingratiating. It is also frustrating and irritating. Its graces result from Dr. Harrower's ability to be charming, to address herself to the concrete minutiae of the problems of clinical practice, and to share her maturity and wisdom as a seasoned clinician with her readers. Its irritants are the products of a highly biased point of view for which the author, to give her due credit, takes sole responsibility. But her text would certainly have been less misleading if she had entitled it *One Woman's Practice of Psychodiagnostic Clinical Psychology in an Eastern Metropolis*.

Dr. Harrower's latest offering represents the winnowing of her many fruitful years of experience in practice. It is her attempt to let her audience know how she has wrestled successfully with the task of supporting herself as an independent practitioner of clinical psychology and to set forth the values she sees as flowing from and supporting that role. In her preface, she commends her text to diverse audiences. At one and the same time she hopes to describe the vicissitudes of clinical practice to those psychologists considering an entrance into its arcane rites, to graduate students, and to members of allied professions. It would unfortunately be a tragedy if any or all of these groups accepted Dr. Harrower's descriptions as the final word on the nature of the profession.

This intensely personal account, many portions of which have literally been dragged from obviously bulging files of papers, case reports, speeches, and informal addresses, is misleading to its readers on several accounts. While some of Dr. Harrower's particular points of view are explicit, others, unfortunately, are not.

First, the author is concerned about the clinical psychologist's identity struggle. She acknowledges that the majority of her colleagues in practice have chosen to serve as psychotherapists. But she devotes only one meagre chapter to the psychologist as therapist, and even most of the content of this section is taken up with the correlative use of diagnostic materials in furthering treatment plans! Her rationale is stated thus: "(My emphasis) is not an oversight, however; rather I have done it deliberately for several reasons. Among them is the fact that I am concerned here with discussing the practice of a new profession. To the extent that we

are psychodiagnosticians and research consultants in practice we, as psychologists, have indeed staked out and developed a new professional entity. But the practice of psychotherapy — psychotherapy unrelated to psychodiagnosis and research — is not new and far from being able to say we were part of its emergence as a societal entity, we have to admit that we are not always, and in all places, even certain of a welcome into the therapeutic fold!" Clearly, Dr. Harrower is saying that she is not going to describe the practice of clinical psychology as it is actually undertaken but is going to describe it as she feels it ought to be. One can only wish that this statement appeared on page one of her book and that the reader did not have to wait until page 135 to discover what the author was up to.

Dr. Harrower not only slights psychotherapy as a legitimate activity for the clinical psychologist, but she entirely omits potentially provocative attention to the practicing clinician as teacher, supervisor, and/or mental health consultant to groups — roles which the reviewer's colleagues actively seek out and find extremely gratifying. And even her preoccupation with psychodiagnosis is heavily loaded with detailed accounts of the use of objective and projective tests in arriving at an overall personality description. One misses attention to and an exposition of the clinician's own sensitivities and awareness of his own subjective reactions to a patient as potentially rich sources of diagnostic insights.

A final theme which many practicing psychologists may find an anathema sounds around the relations of clinical psychology to medicine. Starting with the statement that the majority of her referrals come from members of the medical profession (the reviewer and his colleagues get less than ten percent in this fashion) and going on to the injunction that the fledgling practitioner should seek out a "mental health team" and have a majority of his patients referred by physicians, Dr. Harrower seems to favor an ancillary status for psychology. She is painfully aware of the controversy which exists about the independent practice of clinical psychology. She, herself, for years has struggled in the vanguard to win recognition for her colleagues. It is possible, however, that due largely to the efforts of dedicated clinicians like the author the issues and battles of the last two decades have reshaped the field and that it is not quite so necessary to be apologetic about the legitimacy of psychology's claims. To use an analogy from another sphere, do we now

need George Washington Carvers or Martin Luther Kings?

To sum up, Dr. Harrower has written an excellent source book for clinical psychologists interested in conducting a psychodiagnostic private practice and for others curious about how some psychologists function in the privacy of their offices. It is filled with sage advice; with practical suggestions on everything from fees, to choice of a test battery, to referral promotions, to contacts with various non-psychological publics; and with the charming, individualistic revelations of one of the grand figures of the profession. Dr. Harrower's energy is boundless. She states, for example, that a full test battery and a detailed report should take no more than four hours to accomplish — an awesome comment on her own performance that few of us could hope to live up to. But this highly personalized account of Dr. Harrower's professional life has serious, almost crippling defects as a guide to current practices or to developing historical trends at large in a profession in ferment.

ARTHUR L. KOVACS
Western Psychological Center
Sherman Oaks, Calif.

Wagner, Edwin E., *The Hand Test*, Akron, Ohio, The Mark James Company, 1962. \$7.50.

Brickin, Barry; Piotrowski, Zygmunt A.; Wagner, Edwin E., *The Hand Test*, Springfield, Illinois, Charles C. Thomas, 1962. \$5.00

About ten years ago, this reviewer became interested in the proposition that the hands, being an important and expressive symbol of communication, might be useful in the study of projected feelings and responses. Accordingly, he conceived a series of expressive movements of the hands whereby they might elicit a system of responses. He was prepared to photograph the hands, singly and in pairs, and then attempt to ask examinees what the hands were doing, what they represented, and what action was taking place. The examinee was to make up stories based upon the series of photographs.

Unfortunately for him, the concept never developed beyond the planning stage and it was with ambivalent feelings that he learned that Edwin Wagner had conceived of the same idea and had done something about it. He was disappointed that someone else had the same idea and pleased that his idea was

vindicated to the point of having a workable and meaningful test developed from it.

Wagner's test consists of a series of ten cards on nine of which a hand has been drawn. The last card is blank, very much like that of card 16 on the TAT. The cards are presented one at a time and the subject is asked to tell what the hands are doing. For the last card, the subject is asked to imagine a hand and tell what it is doing. Responses are recorded verbatim along with initial response times per card. The responses are then scored and interpreted in accordance with a somewhat formal and a bit involved procedure. The test, to some extent, represents a cross between the Rorschach in its scoring, timing, observation of card turning, and interpretation and the TAT in the form of its responses and the possibility of analysis and interpretation without some of the need for formal scoring. Too, the scoring resembles some factors of those of Murray and Tomkins in that we find scoring compartmentalization into affection, dependence, communication, exhibition, direction, aggression, acquisition, active, passive, tension, crippled, fear, description, bizarre, and failure. The author states that the test can be administered in about ten minutes, scored in about five, that it is completely non-threatening and can be easily administered to depressed, deteriorated and hostile subjects. Norms, based upon more than 1,000 protocols, from six years of age and up are included. Reliability was based on the independent scoring of three graduate student scorers of 100 protocols and ranged from .86 to .96. The author recognizes that in the development of the protocols for the total 1,020 cases the subjects mostly resided in Ohio at the time of testing, that the N's are low in some categories, and that the seventeen groups of scoring categories are only the major but not the total of those possible. The drawings are fairly clear but some may lend themselves to misinterpretation (possibly pictures of hands would have been clearer), the manual is clear and complete, and the one sheet scoring blank which has on one side space for recording the initial response time, the responses, and the scoring for each card, and on the reverse side the summary sheet containing the name, address, and other identifying information as well as the ratios, qualitative and administrative observations, case history and diagnostic data, and diagnosis appears to be well-developed. Even though administrative and scoring time would have been extended, it might have been useful to have included some additional cards in which two

hands were in some form of relationship. It is hoped that the author will experiment with this to determine whether further development of affectional relationships could be elicited.

The test was originally designed to predict overt aggressive behavior and the research monograph by Bricklin, Piotrowski, and Wagner entitled "The Hand Test" reports on the study involved in this endeavor. It consists of a description of the test, scoring system, study of the special groups investigated, and clinical applications. The authors point out that the primary goal was to determine whether the test responses would differentiate among groups of whom information was possessed concerning aggressive acting-out tendencies. Their findings indicated that prison inmates and hospital acting-out cases earned significantly higher acting-out scores than did indigents, non-acting-out hospital cases, and normals. Comparison is made between the Hand Test and the Rorschach with the authors pointing out that responses to the Hand Test are less imaginative and that more of the subject's fantasy life can be inferred from the Rorschach.

The authors of the book contend that the Hand Test serves a useful purpose in clinical studies of juvenile delinquents, children with reading problems, schizophrenics, organic patients, and epileptic patients. They consider the test another useful tool in clinical practice to be used in conjunction with and not in place of such other devices as the Rorschach and Thematic Apperception Test. This writer concurs in this view and believes that its true value may be realized after further use and research studies. He would like to see whether photographs rather than drawings would give greater results and whether the use of additional cards using two hands would elicit further meaningful responses. The book should be used in addition to the test and its manual because of the additional protocols and the extensive case summaries and interpretations.

IRVING R. STONE
State Mental Hygiene Clinic
San Diego, California

Witkin, H. A., Dyk, Ruth B., Faterson, Hanna F., Goodenough, D. R.; & Karp, S.A. *Psychological Differentiation*. New York; John Wiley & Sons, 1962. xii, pp. 418, \$7.95

extended series of programmatic investigations carried out under the direction of the principal author, who has long been interested in the influence of personality upon perception. The major observation from which this work stems is the empirical one that certain individuals are perceptually field dependent; i.e., are susceptible to the influence of the environmental surround while other individuals are more or less independent of this influence. On the basis of these findings, together with the impression that these field dependent or field independent characteristics seem relatively stable within a given individual, the differentiation hypothesis has been advanced. In essence, this hypothesis states that individual development proceeds from a state of minimal differentiation toward a state of optimal differentiation. In this context, differentiation is considered to be reflected by an increasing articulation or structuring of the experiential world such that the self becomes increasingly differentiated from the environment with a concomitant development of the awareness of the self as a discrete entity. From this position certain hypotheses are proposed involving articulation of the experiential world, articulation of the self as reflected in body concept and a sense of separate identity, and articulation of controls and defenses with regard to the self.

The research here reported addresses itself to these questions systematically and in some detail, ranging from the differences observed between the field dependent and the field independent individuals on perceptual variables to the differences observed between the mothers of the field dependent individuals and the mothers of the field independent individuals. A variety of techniques have been brought to bear upon the topics investigated, including both laboratory studies of the perceptual process with standard experimental techniques and clinical studies of the perceptual process as reflected in the Rorschach and the Thematic Apperception Test.

The volume can roughly be considered in three portions. The first portion of the work deals with individual differences in differentiation; i.e., research demonstrating that there are differences in perceptual approach among individuals and that these can be conceptualized as reflecting a continuum from field dependent experiencing of the environment to field independent experiencing of the environment. The second portion of the book deals with an attempt to explore the differentiation hypothesis as it is reflected in personal functioning in such areas as general

The present volume is the resultant of an

activity, verbalization, or psychopathology. The third and final portion of the book explores the possible origin of these differences in differentiation as stemming from maternal and other influences.

The material presented hews to the empirical line and is presented straightforwardly. Speculation is labeled as such, and, although not entirely absent, is not often present. A circumspect approach to speculation is certainly commendable but heuristic speculation, done well, is a minor vice and should be indulged on occasion.

The authors indicate that a great deal of data remains to be analyzed, especially that obtained in collaboration with Lois Murphy from her researches at the Menninger Foundation. It is tantalizing to speculate upon these potential findings and to wonder what determinants of differentiation can be observed at early ages. However, the decision to present available data now, rather than later, is a wise one. Often, if reports of de-

velopmental research spanning long intervals are not periodically presented, the sheer magnitude of the problems involved in organization of the data may deter the investigator from an adequate presentation of his findings.

This volume is rich in empirical findings. No review, within the scope of allotted space, can encompass the diversity of research reported in the present volume. Further, this work is well worth reading in its entirety as an exemplification of what programmatic research planning can achieve by way of systematically exploring an area of investigation. It presents a number of intriguing research possibilities, both clinical and laboratory. This volume will be of particular interest to the developmental theorist, the personality theorist, and the clinician. I commend it to the reader.

A. BARCLAY, PH.D.
Saint Louis University and
Cardinal Glennon Memorial
Hospital for Children

ANNOUNCEMENTS

The following membership changes have been approved by the Board of trustees.

<i>Fellows:</i>	Charles R. Dewitt
Seymour Blumenthal	John R. Donoghue
Charles O. Gaston	Cono Galliani
Jack L. Herman	Tom W. Patterson
Burdette Lundy	<i>Student Affiliates:</i>
Robert McCully	Curtis Dye
Richard Siegal	Alan S. Mishne
	Enid R. Williams

Associates:

Elianore I. Brassard

TRAINING IN THE RORSCHACH METHOD

General training in the Rorschach Method and other projective techniques by the Tavistock Institute and Clinic was initiated in 1947. This was succeeded by the development of a three-year course for those living in or near London, who could therefore attend weekly seminars. In 1954 requests from those living far from London or in other countries, led to the establishment of summer schools, in which the work was intensive in character so that considerable teaching could be carried out in short periods. Details of the three-year course may be found in the Tavistock general prospectus of training. Information regarding summer schools is given below.

Summer Schools Programme: August 1963

Consultant and Senior Tutor:

Theodora Alcock, F.B.Ps.S.,
Vice President, British Rorschach Forum,
Fellow, Society for Projective Techniques,
U.S.A.

With the assistance of colleagues.

(a) *Introductory* 12th/16th August 1963

Lectures on theory, administration, scoring and principles of interpretation. Case demonstrations and group discussions. Study of normal and pathological Rorschach material.

(b) *Intermediate* 19th/23rd August 1963

Continuation of introductory course, with more emphasis on principles and practice of interpretation and on validation of data.

(c) *Advanced* 26th/30th August 1963

Lectures and group discussions on advanced interpretation, and differential diagnosis. Interpretative study of case material of various types, using the psychogram, together with sequence analysis. Problems of scoring and of reporting to colleagues, etc.

The groups at each level are confined to a small number working for 6 hours every day from Monday to Friday inclusive, beginning each day at 10 a.m.

Members of all groups are invited to send for discussion their own case records, these being if possible typed.

Early application is requested:

The Secretary,

Tavistock Institute of Human Relations,
3 Devonshire Street, London, W.1.

Note: Professional qualification in psychiatry or psychology is a pre-requisite qualification for membership of all courses.

INDEX

CONTENTS OF VOLUME 26 (1962)

ARTICLES

	ISSUE	PAGES
ABRAMS, S. A refutation of Erksen's sensitization-defense hypotheses.....	3	259-265
APPELBAUM, S. A. and HOLZMAN, P. S. Answer to Dr. Jerome D. Pauker's note on "The color-shading response and suicide".....	4	430
APPELBAUM, S. A. and HOLZMAN, P. S. The color-shading response and suicide.....	2	155-161
BERGER, L. and EVERSTINE, L. Test-retest reliability of the Blacky Pictures test.....	2	225-226
BLOCK, W. E. Psychometric aspects of the Rorschach technique.....	2	162-172
BLOOM, B. L. The Rorschach popular response among Hawaiian schizophrenics	2	173-181
BLUM, G. S. A guide for research use of the Blacky pictures.....	1	3-29
BRUCK, M. A note on modified instructions for Bender-Gestalt elaborations and associations.....	2	227
CARP, F. M. Psychosexual development of stutterers.....	4	388-391
CHAGNON, G. (See Fuller, J. B.)		
CHRISTY, E. G. (See Simmons, W. L.)		
COLEMAN, J. C. (See Turner, G. C.)		
COLLIER, H. L. (See Murstein, B. I.)		
COROTTO, L. V. and CURNUTT, R. H. Ego strength: a function of the measuring instrument.....	2	228-230
CRUMPTON, E. Projective case study of a true hermaphrodite.....	3	266-275
CURNUTT, R. H. (See Corotto, L. V.)		
DANA, R. H. The Validation of projective tests.....	2	182-186
DAVID, C. Interpersonal measurement of two occupational interests groups.....	3	276-282
EHLINGER, H. (See Knapp, R. H.)		
EHRENREICH, G. A. (See Siegal, R. S.)		
EVERSTINE, L. (See Berger, L.)		
EXNER, J. E., Jr. A comparison of the human figure drawings of psychoneurotics, character disturbances, normals, and subjects experiencing experimentally induced fear.....	4	392-397
EXNER, J. E., Jr. The effect of color on productivity in cards VIII, IX, X of the Rorschach.....	1	30-33
FILMER-BENNETT, G. and KLOPPER, W. G. Levels of awareness in projective tests.....	1	34-35
FISHER, S. Relationship of Rorschach human percepts to projective descriptions with self reference.....	2	231-233
FLUCKIGER, F. A. (See Weinberg, G. H.)		
FOSTER, H. C. (See Weisskopf-Joelson, E.)		
FULLER, J. B. and CHAGNON, G. Factors influencing rotation in the Bender-Gestalt performance of children.....	1	36-46
GEWIRTZ, H. (See Knapp, R. H.)		
GLUCK, M. R. (See Leventhal, T.)		
GOLDFRIED, M. R. Some normative data on Rorschach developmental level "card pull" in a psychiatric population.....	3	283-287
GOLDFRIED, M. R. Rorschach developmental level and the MMPI as measures of severity of psychological disturbance.....	2	187-192
GOLDMAN, A. E. Symbol consensus and univocality in schizophrenia.....	3	288-294
HAWORTH, M. R. Responses of children to a group projective film and to the Rorschach, CAT, Despert Fables and D-A-P.....	1	47-60
HERSCH, C. The cognitive functioning of the creative person: a developmental analysis.....	2	193-200

	ISSUE	PAGES
HOLZBERG, J. D. (See Knapp, R. H.)		
HOLZMAN, P. S. (See Appelbaum, S. A.)		
KLOPFER, W. G. The role of diagnostic evaluation in clinical psychology	3	395-398
KLOPFER, W. G. (See Filmer-Bennett, G.)		
KNAPP, R. H. and EHLINGER, H. Stylistic consistency among aesthetic preferences	1	61-65
KNAPP, R. H., GEWIRTZ, H. and HOLZBERG, J. D. Some personality correlates of styles of interpersonal thought	4	398-403
KORNER, A. F. Developmental-diagnostic dimensions as seen through psychological tests	2	201-211
LAPKIN, B. (See Silverman, L. H.)		
LEVENTHAL, T., GLUCK, M. R., SLEPIAN, H. J. and ROSENBLATT, B. P. The utilization of the psychologist-patient relationship in diagnostic testing	1	66-79
LEVINE, M. and SPIVACK, G. Human movement responses and verbal expression in the Rorschach test	3	299-304
LEVITT, E. E., LUBIN, B. and ZUCKERMAN, MARVIN A simplified method of scoring Rorschach content for dependency	2	234-236
LUBIN, B. (See Levitt, E. E.)		
LUSTGARTEN, B. J. (See Palmer, J. O.)		
MCCULLY, R. S. Certain theoretical considerations in relation to borderline schizophrenia and the Rorschach	4	404-418
MCGREEVEY, J. C. Interlevel disparity and predictive efficiency	1	80-87
MCGUIRL, D. and MOSS, C. S. An indirect validation study of the Draw-A- Person test through the cartoons of William Steig	1	88-95
MARKS, P. A. (See Seeman, W.)		
MAYMAN, M. Post-doctoral training for clinical competence	3	305-309
MOOS, R. H. Effects of training on students' test interpretations	3	310-317
MOSS, C. S. (See McGuirl, D.)		
MURSTEIN, B. I. and COLLIER, H. L. The role of the TAT in the measurement of achievement as a function of expectancy	1	96-101
NEIGER, S., SLEMON, A. G. and QUIRK, D. A. The performance of "chronic schizophrenic" patients on Piotrowski's Rorschach Sign List for organic CNS pathology	4	419-428
NEIGER, S. (See Quirk, D. A.)		
NEURINGER, C. Manifestations of anxiety on the Rorschach test	3	318-326
NICHOLS, R. C. (See Strumpfer, D. J. W.)		
PALMER, J. O. and LUSTGARTEN, B. J. The prediction of TAT structure as a test of Rorschach's experience-balance	2	212-220
PAUKER, J. D. A simple method for measuring the area of a figure drawing	2	237-238
PAUKER, J. D. Base rates in the prediction of suicide: a note on Appelbaum's and Holzman's "The color-shading response and suicide"	4	429
PHELAN, J. D. Projective techniques in the selection of management personnel	1	102-104
QUARRINGTON, M. (See Quirk, D. A.)		
QUIRK, D. A., QUARRINGTON, M., NEIGER, S. and SLEMON, A. G. The performance of acute psychotic patients on the index of pathological thinking and on selected signs of idiosyncrasy on the Rorschach	4	431-441
QUIRK, D. A. (See Neiger, S.)		
ROSEN, I. C. (See Siegal, R. S.)		
ROSENBAUM, I. S. (See Silverman, L. H.)		
ROSENBLATT, B. P. (See Leventhal, T.)		
ROSENTHAL, M. Some behavioral correlates of the Rorschach experience-balance	4	442-446
ROTHAUS, P. Problems in the measurement of aggression-anxiety	3	327-331
SACKS, H. The effect of preliminary verbal conditioning on inkblot test responses	3	332-336
SCHAE, K. W. The performance of mentally defective children on the color pyramid test	4	447-454
SCHULBERG, H. C. and TOLOR, A. The "meaning" of the Bender-Gestalt test designs to psychiatric patients	4	455-461
SCHUMAN, E. P. A scoring rationale for the Sargent test of insight into human motives	4	462-468
SEEMAN, W. and MARKS, P. A. A study of some "test dimensions" conceptions	4	469-473